Novel approaches to biometric security with an emphasis on liveness and coercion detection

Edge Hill University

Peter William Matthew

Department of Computing

Edge Hill University

This dissertation is submitted for the degree of

Doctor of Philosophy

January 2016

I would like to dedicate this thesis to my loving and supporting parents...

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text. This dissertation contains less than 80,000 words including appendices, bibliography, footnotes, tables and equations and has less than 150 figures.

Word Count: 72,079

Peter William Matthew
January 2016

Acknowledgements

Throughout the past five years, there have been countless people who have offered advice, friendship and fascinating discussion. To them, and everyone else, who has listened to me drone on about my research boring them with minute details on a subject they mostly don't care about, you have my thanks, my appreciation and my unending gratitude.

I would like to thank my supervisors, Professor Mark Anderson and Dr Maybin Muyeba for all of the support and guidance they have provided. Dr Muyeba has provided fantastic insight and feedback over the whole process, for which I am very grateful. I especially would like to thank Professor Anderson for his constant patience, good spirit and ability to motivate, especially during the darker days of this research. Mark has always been willing to discuss my research, provide fantastic ideas, inspire with motivating words, and respond to the incessant emails with feedback despite the time of day or night. His example has inspired me to strive to go as far as I can and for this I will be eternally grateful.

Secondly I would like to thank my examiners, Dr Chris Beaumont and Dr Siraj Shaikh for taking the time to examine my Ph.D. and for all of the fantastic feedback provided. You have both helped me enormously, and I am eternally grateful to you both for making the culmination of this research run as smoothly as possible.

I would like to thank the Computing Department at Edge Hill University; there have been countless times, over these years that I have enjoyed chatting about my research and the life of a graduate researcher, with many people offering good advice. I would especially like to thank Collette, Dave and Mark Liptrott for their advice and support during this time.

I would especially like to thank Besim Mustapha for his advice, discussions and the chance to conduct collaborative research, it has been fun and a highly positive experience I will always remember. I would be remiss without thanking Dan Kay, the computing technician, for all of the help and by having good nature in the face of constant requests for software and hardware, thanks, Dan!.

I would like to thank the members of the Language Centre, and the teaching staff on the IFP and PMP programmes, especially Linda and Kate who have both listened to my constant, often boring, talk about my research and who have offered countless invaluable insights.

After being the only Ph.D. student in the department for the first year of my research, I have a complete appreciation of the support structure and comradeship that the other Ph.D. candidates have provided. Dan, Darryl and Alex have always been available to have a chat about research, and more importantly to be able to bounce ideas off. I wish them all the best in their research.

I would like to thank my parents for all of the support they have given me, and the constant acceptance of the everlasting student. They have done everything in their power to make my research time pass as easily as possible, and they have always been there when times have been difficult. They have always provided an incomparable pillar for me to hold on to, and I will never be able to thank them enough for their support and love, and I will always be immensely proud of them both.

Throughout this entire process, I have had a group of fantastic friends supporting me, friends that have been together for many years. I would like to thank Greg and John for their chats and insights as they have provided great ideas. I would like to thank Alan for his time, willingness to discuss my work, and offers of support over the years, something that I will always appreciate. Finally, I would like to thank Bill for his unending support and help over these past years. Bill has always been there to take me out of the work, offering support, checking to see if I was all right and general watching out for me, without his support the overall outcome might have been quite different. A chap could not ask for a better group of friends.

The final thanks are arguably the most important one. I would like to thank my dog, Robin, for his constant companionship, love and the joy he has brought to my every day, especially during the late night sessions where he, in his basket, has kept me company throughout this past age.

To anyone I have missed I apologise profusely and give my heartfelt thanks.

Table of contents

List of figures

List of tables

Chapter 1

Beginning to consider

+++Mr. Jelly! Mr. Jelly!+++
+++Error At Address: 14, Treacle Mine Road, Ankh-Morpork+++
+++MELON MELON MELON+++
+++Divide By Cucumber Error. Please Reinstall Universe And Reboot+++
+++Whoops! Here Comes The Cheese!+++
+++Oneoneoneoneoneoneone+++

Pratchett, 1994

Conducting any research demands that a structured plan is in place and followed to maintain purpose and provide structure. This research is no different and there is a variety of aims and objectives that have been identified that promote this process and identify how the route will be formed. The overall aim of this research is *'To develop a novel taxonomy that enables an in-depth analysis of different liveness detection techniques that can be applied to other novel areas such as the development of coercion detection technologies culminating in a predictive and analytical method'*. The first stage of this path is to identify what are the constituent areas within this research.

1.1 Context

To live in a world where the operations of daily living are automated by a home, work, or a public system has been the vision of many for some years [178]. Amongst the primary

concerns is the comparative complexity of systems and the ever-increasing security considerations [156]. This emphasis on security is a standard expectation within the current computing environment, however, to better represent the complexity a more dynamic security technique is being considered. Biometric security techniques offer a reliable and secure alternative compared to physical and key based security; systems such as passwords and pins. The use of biometric systems opens up a host of other considerations and possibilities such as the integration of health and well-being monitoring, as well as providing a more data-focused society with information about lifestyles and habits as identified within [55]'s work on smart homes and biometric devices.

Traditionally biometric security has been rarely used, partially because due to the lack of user trust and willing organisations to implement the technology. However, the amount of research within the area of biometric security has increased massively over the past ten years due to the general increase in implementation. These implementations range from physical to logical access and can be used in either identity and surveillance scenarios. It is the area of logical access that is growing the most, access PC, smart devices, etc., and this is projected to have a 10% increase in revenue going from 21% to 31% in the upcoming two years [30]. Along with the 1% growth of physical access systems, this proves that the market and importance of biometric solutions is increase and will affect the security concerns of companies in the upcoming decade. This can be seen by the forecast for the next five years which predicts that the annual revenue will increase from, in millions, \$1.621 to \$34.637 [109]. Therefore, the impact of these security techniques will become increasingly important and will lead to emphasis on efficient and effective threat detection

Circumvention of biometric techniques using spoof samples became the main problem to contend with when developing biometric security solutions. Liveness detection was deemed to be the solution, by making it harder to accept spoof samples, and this led to an explosion of liveness detection research into both solutions and causes. Due to this sudden influx of research, there was minimal focus placed on creating techniques that followed similar processes or produced similar outputs. This caused confusion as many techniques followed similar lines of thought, however, did not follow a similar process, therefore, minimising the impact they could have had. This was also seen in the use of different measurements often being relevant only to the research they were contained in and other techniques. If there had been a process to align these research projects, then they could have been easily compared to each other, and the most suitable would have been highlighted. This allowed a huge amount of liveness techniques to be created that was very difficult to compare, which prevented efficient implementation into biometric systems both current and future. This has led to confusion when choosing liveness detection techniques for implementation both in

multi-modal and fusion environment, and has raised the question should some method of categorisation be sought to ease the comparison between liveness techniques? Therefore, promoting only the most suitable techniques for the specific installation. This situation could also potentially reoccur for coercion detection research, as there is currently a minimal amount in this area. However, if the trend from liveness detection is followed, then the amount of research will reach critical mass rapidly once again creating a situation that prevents the ease of comparison and implementation. Therefore alongside the development for liveness detection, it is the intent not to allow this liveness research bloat to occur again and instead, create a method of categorization that can be used for both liveness and coercion detection techniques. To help categorise these techniques, a taxonomy will be developed that will highlight the salient features of these techniques, therefore, promoting a more efficient analysis and comparison intra-technique. There is a precedent in this area, as generic biometric security techniques contain a basic taxonomy. However this research will look at the practicalities of this taxonomy and highlight if it can be used for liveness detection, and if so what, if anything, will need to change.

While these concepts are relevant they are fraught with flaws and potential problems. The integration of biometric techniques alone comprises of numerous factors that superceded the considerations of traditional security systems such as the problems surrounding spoof data, and the reliance on user capabilities (resistance to phishing, etc.), therefore the inclusion of autonomic environments can potentially apply a more responsive and accessible system to all users. These factors also bring up numerous questions such as how to identify the best technique for installation, what other security considerations are important within biometrics such as liveness detection and what future novel approaches need to be identified.

1.2 Aims

Security has always been one of the paramount factors within any computing system and it is becoming increasing difficult to ignore the added security concerns that a ubiquitous deployment of biometric security has identified, such as the 'collectability' of samples for spoofing. Subsequently, both traditional and specific threats must be considered when conducting biometric research.

Therefore, it is the aim of this research to identify the specific factors that liveness and coercion detection requires, coupled with the most suitable methods of implementation that can be derived from both. This is not the sole aim of this research; the second aspect is to identify the most suitable method of system implementation specifically considering the applicability of autonomic environments. This is because of the huge complexity of

integrated biometric environments and the innate advantages autonomic environments can provide potentially linking into smart environments.

These aims, when achieved, will culminate in the creation, development and deployment of a taxonomy that will classify different biometric security techniques and allow a quantitative classification to be provided showing the relative increase or decrease of security threat within a system, measured by the algorithm developed throughout this research.

Furthermore, it has been identified within current research that the amount of standardisation and cross technique, understanding within liveness detection is very poor and would benefit from a more thorough standardisation. Additionally, as coercion detection is in its infancy the development of different coercion detection techniques and classifiers is groundbreaking in scope.

Therefore, the intent of this refined aim is to identify, categorise and evaluate different biometric liveness detection techniques alongside the development and classification of novel coercion detection techniques that will enable future techniques to be implemented. To refine the overall aim into a specific statement, the following identifies the key aims, and the overall purpose of the research.

Therefore, the purpose of the research is twofold firstly: to develop a novel taxonomy that enables an in-depth analysis of different liveness detection techniques, and that can be applied to other novel areas such as the development of coercion detection technologies culminating in a predictive and analytical technique. Secondly, while the overall aim is comprised of some factors the individual factors therein comprise of the objectives that will be sought throughout the research process. Therefore, the following objectives will provide practical guidance that will allow the overall aim to be sought and accomplished.

1.3 Objectives

In achieving the aim the following objectives are highlighted:

1. Critically evaluate, in sufficient detail, the different algorithms, models, architectures and associated technologies within the biometric environment. This will be done within chapter three and be based on the following two areas:

   (a) Biometric Security

   (b) Liveness detection

2. Develop a taxonomy for security techniques that can cover a range of different current and future technologies. Focusing on liveness detection, but maintaining scalability

allowing future techniques to be adapted for the taxonomy while creating output that can be compared to other techniques, therefore creating a comparative and analytical tool. This will be done within chapter four.

3. Investigate how the inclusion of different levels of liveness and coercion detection affects the security of the system. This will be done within chapters five and six.

4. Identify the underlying structure of coercion detection and how it should work with biometric security, linking to both liveness detection and biometric authentication. This will be done within chapters five and six.

5. Develop a baseline coercion detection environment, to determine the relevant validity of coercion detection in current biometric systems as well as how to measure coercion detection. This will be done within chapters five and six.

6. Evaluate the taxonomy by testing it with a range of data gathered from respected peer reviewed research. The validity of such data is important and has been discussed within chapter 2. This will be done within chapter and six.

   (a) Identify suitable research that can be used for data collection

   (b) Apply this data to the taxonomy

   (c) Evaluate the results

1.4 Original Contribution to Knowledge

The applicability and originality of research is a huge factor and one that has been closely considered during the development process of this research. Therefore, the specific original contribution to knowledge has been identified in summarised form for ease of access.

1. Develop an interval liveness detection based taxonomy that outputs relevant numerical standards that in turn can be used within a variety of situations. This is not something that exists currently as the only classification system uses an ordinal system that provides very little overall comparative data. This is highlighted within chapters three and four.

2. The taxonomy will highlight very specific points of interest within different techniques, allowing a detailed dissemination and analysis of the overall technique. This is something that has not been attempted currently and will be highlighted within chapter five and six.

3. The taxonomy will be able to predict security levels; this will be done by entering associated data from similar case studies allowing developers more idea of how their implementation would function. This is highlighted within chapters four and five.

4. The current focus on biometric security is centred around liveness detection standards as they have been the most important sub-area within biometric security. As the technology has improved over the years and liveness has become more accepted, a further concept has started to become more and more relevant. This concept is coercion detection and would occur after the liveness detection techniques have acted. Currently, there is very little information in this area despite its potential security flaws. This is highlighted within chapters three, four and five.

5. The development of the taxonomy that can be adapted also to function with the same, or similar, classifiers within coercion detection. This is so that the planning and development will consider the application in different areas instead of trying to adapt the final taxonomy to fit different areas. This is a move to create an efficient and robust taxonomy. This is highlighted within chapter five.

6. Due to the relative youth of coercion detection there are very few techniques, therefore, by developing and presenting a variety of novel techniques, new ground is being covered. This is highlighted within chapter five.

1.5 Thesis Structure

Chapter One contains the introduction, sets the context and aims of the research and highlights the original contribution to knowledge along with publications gained from this research. *Chapter Two* will cover the methodological choice, which is grounded theory, as well as highlighting other potential methodologies that while are viable have not been used.*Chapter Three* will identify the background to biometric security while identifying some of the key areas that are currently lacking such as an appropriate way of measuring liveness detection techniques as well as the entire coercion detection sub-discipline. *Chapter Four* discusses the development of a new taxonomy that will classify liveness detection while moving away from the current ordinal measurement system used within the research area. Analysis of these liveness classifiers will then follow leading into the adaptation of the taxonomy of coercion detection techniques. Further development of these new techniques will follow, identifying metrics for coercion detection and an analysis of the proposed classifiers. After the taxonomy development *Chapter Five* analyses coercion and liveness techniques by

applying the taxonomy across a selection of liveness and coercion techniques. This will then be followed by the development of an algorithm to denote the level of security an individual technique has achieved. Explanation of the algorithm development, components and testing will then be included. Finally *Chapter Six* will contain the final concluding remarks and will cover some of the areas in the future that can be looked into, alongside some focuses for article and conference submission.

1.6 Published Work

As part of this work numerous papers were submitted and conferences attended, culminating in the following:

1 P. W. Matthew and M. R. Anderson, "Novel Approaches to Developing Multimodal Biometric Systems with Autonomic Liveness Detection Characteristics," in *Intelligent Systems for Science and Information*, vol. 542, L. Chen, S. Kapoor, and R. Bhatia, Eds. Cham: Springer International Publishing, 2014, pp. 147– 159.

2 P. W. Matthew and M. R. Anderson, "Biometric Incorporation in Pervasive and Autonomous Systems Emphasising the use within e-Health Specific Smart Homes," *Int. J. Intell. Comput. Res.*, vol. 2, no. 1/2/3/4, pp. 211–218, 2011.

3 P. Matthew, "Autonomous synergy with biometric security and liveness detection," in *Science and Information Conference (SAI)*, 2013, 2013, pp. 376–382.

4 P. W. Matthew, "Biometric implementation in autonomous systems with an emphasis on smart home applicability," *International Conference on Information Society (i-Society 2011)*. pp. 382–387, 2011.

5 P. Matthew and M. Anderson, "Novel Categorisation Techniques for Liveness Detection," in *2014 Eighth International Conference on Next Generation Mobile Apps, Services and Technologies*, 2014, pp. 153–158.

6 B. Mustafa, P. Matthew, and F. Naveed, "Leveraging smart monitoring and home security technology for assisted living support," in *Science and Information Conference (SAI)*, 2013, 2013, pp. 114–123.

7 B. Mustafa, P. Matthew, and F. Naveed, "A Smart Monitoring System for Assisted Living Support Using Adaptive Lifestyle Pattern Analysis," in *Intelligent Systems for Science and Information*, L. Chen, S. Kapoor, and R. Bhatia, Eds. Springer International Publishing, 2014, pp. 1–24.

Chapter 2

Methodologies

> What a place! What a situation! What kind of man would put a known criminal in charge of a major branch of government? Apart from, say, the average voter.
>
> Pratchett 2004

Over the past ten years, there has been a surge of research into biometric devices and which characteristics can be used as biometrics. Due to this constant innovation and research many new technologies and characteristics have been and are being, considered for biometric identification such as body odour and the composition of exhalation [104]. To highlight these novel techniques, the definition of biometrics was considered, and the identification used is: *"the automated use of physiological or behavioural characteristics to determine or verify identity"* [67]. As more research intimated that further key areas could be applied to biometrics, [114] increased this definition to include chemical characteristics. However before this is considered that techniques associated with research must be considered. Therefore, this chapter will cover the methodological choices that have been used. This will include the overall methodology, discussion on the pertinent factors regrading data collection, ethical considerations and risks that may effect his researcher. The overall aim of this chapter is to justify what methodological choices have been made and to show the suitability for this research.

2.1 Research Question

Due to a boom of liveness research, the area has become over saturated with similar research, which does not conform to a uniformed and efficient standard. How would the development of such a uniformed categorisation standard effect liveness detection research?

Identifying that this boom of research, without a method of comparing it efficiently, can cause confusion, how can this problem be minimised for future threat prevent techniques such as coercion detection.

2.2 Appropriateness of the Research Design

After the research questions have been identified, the first consideration when undertaking any research, is what methodology should be used. Within computer science, there is a variety of methodologies that can be applied depending on a host of factors such as what is being presented, proved, questioned, etc. Some of the methodologies are relevant while others are less so, for example, a build methodology would not be viable for this research as there is no building of software, system, etc. Grounded Theory, while unconventional can be used to identify areas of new interest, but requires valid case studies to provided specificity. Experimental methodologies are widely used but can be limited without an initial theory [14]. The following section will highlight the chosen methodology that is based on Grounded Theory.

2.3 Methodology

2.3.1 Philosophy

There are many different research philosophies available to researchers and choosing the correct one is a difficult endeavour. The first stage was to consider some of the core considerations within research philosophies, what is knowledge and how to interpret it. The two areas that will be focused on are ontology and epistemology. Ontology identifies the nature of things; this is how the research is viewed and what assumptions there can be. While epistemology is the understanding of what is known to be true and is it possible to gain knowledge identifying the nature and limitations of inquiry [76]. Within this research, there will be a realism focus on ontology and a positivism focus on epistemology.

This realism ontology approach holds that there are particulars and universals that allow the classification of abstract concepts as well as a direct method of interpreting the subject-

predicate structure of language and speech [27]. As the main route of this research is to categorise areas that do not currently exist then it makes sense that the realism ontology is used. For example, if "affect" is identified as a universal then the particular could be "fear".

The epistemology choices would follow a positivist approach as the focus would be to the user a grounded theory like methodology to identify areas of interest. This would objectivity identify what occurred following the grounded theory style of open-mindedness, and would then be empirically tested and logically analysed. The understanding here is that there are areas to discover and with an open mind and range of experiences theories can be developed, unlike constructivist techniques that identify that knowledge is constructed by scientists.

These factors have affected the choice of methodology for this research so that the ontology philosophy is kept as realism, and the epistemology is kept as positivism. Therefore grounded theory has been chosen as a basis in which to build upon. This will not be the sole techniques considered, nor will it be true grounded theory. Instead, it will make use of the core tenets of grounded theory and will be discussed in detail within the next section.

2.3.2 Grounded Theory

There are many different research methodologies available within computing, each with their advantages and disadvantages depending on the style of research being conducted. This research has been conducted using a methodology based on grounded theory, and whilst this is mostly found in social sciences [23].66, there are a variety of advantages to be had when used within computing. Grounded Theory like methodologies have been used to great effect in computer science areas such as software development [9] and intelligent computing [159]. Grounded theory excels when there is little current theory in an area, but there are some basic concepts present [159]. The main focus of this methodology was to improve the conceptualisation and the abstraction of a theory using logical progression, dedication and analysis[63]. This is especially useful in this research as, whilst there are valid theories and concepts available for liveness detection, the area of coercion detection is woefully under-represented and whilst there are some basic concepts they are not well researched even though they have been identified as important by leading researchers such as [17] and [116].

Grounded theory is based on techniques of experiences. This means instead of beginning with a theory and progressing with its development the current state of a subject is considered and the experiences of the researcher and other associated researchers are consulted to highlight where a theory can be developed [180]. This is especially pertinent in this research as areas such as coercion detection and the taxonomical development features are identified via research experiences as areas that need to be considered in detail. The original work by [63] details the different factors within grounded theory, such as the acceptance of

personal and profession perceptions, as otherwise the integrity of the research would be compromised: as the rigid set of rules highlighted by other research techniques such as the triangulation approach to data gathering in action research [187] or the practical application of problem/enquiry based research [145] would hamper the theory generating grounded theory technique. As the areas discussed within liveness and coercion detection are prone to human behavioural deviations this makes grounded theory, one of the more suitable methodologies to utilise as it allows a variety of factors to be taken into account such as the choices of users, and their reaction to certain stimuli, as well as the researched reaction of the participants.

One of the defining features of grounded theory is the expectation that the researcher will begin the research with an unfettered viewpoint, but also some initial insight into the area being researched and as more information is gathered, as the theories are worked on, the more developed the theory becomes [63] [180]. This is then followed by a selection of theoretical samplings which highlight actions, events, procedures, etc. that can affect the development of the overall theory. However, one disadvantage of using grounded theory is that it normally gathers data by the interview technique which is not viable in this research due to the innate difficulties in gathering primary data related to coercion and liveness [17] [89] such as medical reluctance and ethical requirements. However, instead of utilising a separate research methodology a hybrid methodology will be used, based in grounded theory whilst including case studies. Therefore, this will allow data to be gathered regarding this subject by basing testing and analysis on case study data. This minimises one of the main problems with grounded theory from emerging, the reliance on micro-analysis coding. Instead of having to work through extensive interview transcripts for information [13], the data gathered will be taken from associated cases within the area of liveness and coercion detection. The overall methodology consisted of hybrid techniques which incorporated grounded theory and case studies. The following section discusses these components in more depth, highlighting the areas the research that they apply to as well as highlighted an alternative approach, experimental methodology.

Case Studies

Case studies have been conducted cross-discipline for many years, as it allows the researcher to consider what factors can be appropriate depending on the situation, organisation, etc. and allows the comparison of multiple real-world examples as well as theoretical concepts [180]. Whilst it is used extensively in areas such as health, it has precedence in computing as indicated by the work conducted by [95] and [142].

The main reason for case study inclusion is to allow the in-depth situation and individual exploration that is a key factor of this research methodology. However, as [180] identifies, it is difficult to generalise from case studies and they instead require careful contextualisation or multiple studies need to be conducted to verify the results. This problem is often highlighted by researchers who consider case studies to be an interpretivist style which is because of the emphasis on qualitative data. However as these case studies are being used to highlight specific datasets, they are being used within a positivist environment, which allows the development of general rules that can apply to other studies and theories as highlighted by [185] [64] and [151]. It is this positivist approach that will be utilised; this is due to the wide range of areas that is applied to and the scientific emphasis it contains.

To gather the most appropriate case studies, current research into both coercion detection and liveness detection was considered. The first stage was to identify what constituted an appropriate case study and what did not [180]. To do this, consideration was given to a number of factors such as the scientific nature of the data and the suitability of the source. For liveness detection, this technique will be much easier to accomplish as there is extensive research on liveness standards such as identified by [116] [19] and [144]. However, this becomes much harder to achieve with coercion standards due to the limited current research in the area. Therefore, a specific plan was identified to gather suitable sources to use as case studies.

The first factor considered when choosing appropriate research was to identify the impact it has had. This was achieved by checking for the citations in other publication therefore showing that the work has achieved impact; this was done using the Google Scholar citation metric, as well as publisher produced information such as the IEEExplore citations data. This enabled the identification of salient work in a particular area such as liveness detection competitions and articles from the International Conference of Biometrics, a seminal journal in the area, such as the work conducted by [183]. Alongside this work published in respected peer-review journals such as the International Journal of Emerging Technology and Advanced Engineering [111] and the Journal of Network and Computer Applications [153], were considered. These respected peer-reviewed journals and conferences allowed a very high degree of relevancy to be courted therefore lending credence to the data gathered for and from the research.

As with many areas one problem is the sheer quantity of articles in the biometric area that are published in lesser known journals or conferences therefore not being as relevant as data sources due to the lack of stringent peer review. Whilst there is a plethora of this data it was found that often there was a significant error in the content or the journal was not peer reviewed and therefore it is for these reasons the utilised research was selected.

Whilst most of the research data was selected using this factor some were selected as they were the seminal work or researchers in the area [166] is used as she was one of the original innovators for liveness detection and is still heavily influencing the work such as within the [183] International Conference of Biometrics.

Whilst this technique is effective for liveness detection, due to the minimal research within the coercion area there is an innate difficulty in gathering primary data. Therefore, to pass this limitation, some liveness data was used as a basis, using the same process as mentioned above, which was then adapted to highlight the different coercion techniques developed within the research [17]. This factor was conducted using an experimental methodology that coupled with the grounded theory encompassed the hybrid methodology. Obviously future research would surround the development of actual coercion data sets that can be used within the coercion taxonomy.

Whilst there have been some problems gathering case studies, by either having too much data and having to clear throughout superfluous data: to the lack of data and having to adapt other data as necessary to be of use and whilst other techniques could be employed such as the experimental methodologies. They would not allow the development of the theory and the adaptation of the data so that testing could occur, especially in the face of coercion testing. This is one of the main advantages of grounded theory and case studies. However to highlight some of the alternatives that could be implemented the following section will discuss some of the features of experimental methodologies as experimental is a popular and established methodology in computing such as the work carried about by [57] and discussed in [85].

2.3.3 Experimental

In the early days of computer science research, the belief was that experimental research was not applicable partially due to the belief that computer science was based more in engineering than traditional sciences [31] and therefore did not prescribe to experimental research methods, with some researchers identifying them as potentially useless or even harmful [165]. However, as computer science has evolved so too has the understanding and acceptance of the experimental methodology [165].

The main advantages of experimental methodologies is that it is possible to control minutely the research process and to corroborate theories. This is the main reason that experimental methodology has not been chosen as the main methodology, as there are very few current theories regarding coercion detection and the lack of taxonomical work in both liveness and coercion detection demands a methodology that better allows the development of a theory, which, in this case, is grounded theory [63]. A second factor comes from a famous

statement by Edsger Dijkstra "*an experiment can only show the presence of bugs (flaws) in a theory, not their absence.*". Therefore whilst the work being proposed is theoretical in nature, it will still contain a number of bugs in its development. However the development of the theory takes precedence to the bug fixing. This is most easily achieved using grounded theory, due to the points made earlier. However, the inclusion of experimental methods will also be used to help identify suitable case studies as described above.

A final reason for not using an experimental methodology, as the main focus, is that there are some algebraic equations and, depending on your discipline, it is impossible to prove an equation throughout experimentation as the equation is always correct if correctly constructed and is not prone to variation [165].

These factors have been the key deciding points culminating in the hybrid methodology being chosen; it is key that the grounded theory with case study methodology is adopted, however, to add more applicability and practicality the experimental methodology is also used in the areas mentioned.

2.4 Qualitative and Quantitative Data

One of the first considerations, after the development of the methodology, is: what data will be used, would it be qualitative or quantitative data. This can change the thrust of research in a substantial way and the choice was reviewed solidly [49]. There will be two forms of justification: doxastic and situational. There are some areas that will be researched that have an outcome that is expected and believed, such as the additional of liveness and coercion techniques will improve the security of a system. This is both rational and within normal parameters and therefore its justification is doxastic [173]. Other areas of research will have potential belief, but no belief in them such as; the public are apprehensive and mistrustful of all coercion techniques due to the limited awareness of them. It still is a rational belief, but there is no actual belief in the concept, therefore, showing a situational justification [173]. The hope of this research is to turn these beliefs into knowledge that will begin as perceptual knowledge as it has been gained through seeing and studying certain things. The concept then is to take this perceptual knowledge and make it into indirect perceptual knowledge for the users of the system. Users could implement the findings in some ways such as:

1. Using the research to achieve suitable biometric fusion
2. Provide researchers with tools to further their own research
3. Create an algorithm that can detect best in slot techniques for different systems.

Therefore, a brief identification of data types gathered will be highlighted and what the data will be used for.

Qualitative data is seen as subjective, but this is not necessarily the case [127]. Concepts are often used within qualitative data, concepts that the reader interprets through an understanding of the subject area, thus the concept of fear or embarrassment while using a biometric device would be presented to the subject, this being especially important when considering coercion techniques that may present, rational or irrational, fears for the use. This is an area where the interpretation of the answer leads to a specific concept that can lead to answers such as; unwillingness for the user to draw attention to themselves [49][69]. Qualitative research is very good for gathering descriptive analysis, which concentrates on the importance of the context, setting and the target audience Qualitative data is not without problems and issues, one such problem is that with too many qualitative questions the chance for the answer to be rambling, difficult to read and difficult to extract any useful data from, will become higher [141]. While the use of qualitative data is often ignored, while using grounded theory case studies some qualitative data will be very usefully to determine suitable coercion detection techniques to develop.

Quantitative data consists numerical data. Normally it is easier to correlate and analyse when compared with qualitative data. Many see quantitative data as impartial and objective [127] and [59] postulates that quantitative research methods allow great summarising opportunities and enables the researchers to continue to conduct the research within different settings with a firm backbone of results to compare to. An example could consist of a rated set of biometric styles that a user views, this could, at first, be conducted with a target audience of students, and then people with disabilities, then children, etc. [108]. An important advantage of quantitative data is that any bias held by the research can be negated as there is little to no interpretation of the results [108]. There are some disadvantages with quantitative research and is one of the main and the most damaging issues is that the results that one garners from quantitative research are often narrow in scope and can at times are superficial, not getting into the main issues of the research which is corroborated by [108]. The bulk of the data gathered from this research will be quantitative as it will consist of statistical values regarding the biometric techniques.

While the data being gathered is important one of the key factors to consider is who will make use of the research, and what impact will it have in the area. Therefore the next section will highlight who will use the research and for what purposes.

2.4.1 Target Audience

Throughout the development of the research, there have always been two main end users that would gain from using this research. The following section will detail these two groups:

Researcher

'Researcher' is a broad category incorporating professionals who will utilise the tools created to further their own research either in part in wholly. This can include:

1. The use of the taxonomy to increase liveness and coercion knowledge. Especially within the area of developing new coercion techniques, in this aspect, the coercion research can act as a base level in which it can be improved on and added to.

2. Creating new liveness and coercion standards that can be applied to the taxonomy to check for theoretical suitability before more in-depth testing needs to be conducted.

3. Creating autonomic security environment using the algorithm to highlight the most suitable techniques for a system.

4. Showing, which techniques are lacking or in need of fixing by using the algorithm.

These factors are the main features this research could be used for by researchers; the main focus is on the supply of tools to a researcher such as the taxonomy and algorithm.

Systems Analysts and Developers

While the main area of impact is for researchers, the impact can also be measured from the developers of systems. Instead of using the research as a way to further their own the tools could be used to highlight the best options within a system development such as:

1. The taxonomy can be used to allow developers to choose the most appropriate biometric devices, liveness or coercion techniques for a system identifying that a particular technique has its advantages and disadvantages and showing where the two overlap improving the potential fusion capabilities of the implementation.

2. Creating security system that dynamically change biometric device and techniques depending on user profile, this autonomous environment can improve the overall security of the system and can be used by developers trying to create a pervasive and autonomic environment such as iHomes/iOffices, etc.

These are the main two groups of people who will make use of this research. However there are also some specific areas that could benefit, that are away from the obvious security potential, this is not an exhaustive list it simply highlights some of the main areas that could benefit from this research.

2.4.2 Research Uses

e-Health

While there are many e-health based application that would combine the use and collection of medical data with the liveness and coercion technique to provide a context-aware profiling system that is also able to monitor health needs. A second example would be to improve user accessibility in personal security systems. This would involve users that have an impairment that can, in some way, be alleviated using the research. The reason this area is focused on is that impairments are something that everyone will have to face at some point in their lives. Some people are born with impairments; others develop them due to illness or accidents, yet others develop them over the natural course of ageing.

An impairment may well be a disability that is formally recognised as such by the relevant government and legislation, in the UK a disability is defined as "someone that has a physical or mental impairment and the impairment has a substantial and long-term adverse effect on their ability to perform normal day-to-day activities" [16]. Alternatively, It can also be possible that an impairment is not formally recognised as a disability; this is because it does not have a substantial effect on the normal day-to-day activities as set down in the [16]. Examples could include someone that has joint problems, hard of hearing, sight impairment, learning difficulties such as dyslexia, or any number of other potential impairments that affect the human condition.

Therefore, the research could be used to provide the most suitable biometric technique to a user understanding their personal profile and what techniques would not be suitable, for example, a hippus dilation test would not be most optimal for a user with a sight impairment. There is also additional potential here for users to allow medical data to be gathered from the biometric devices, therefore allowing for tests/monitoring, etc., obviously only with complete ethical and legal backing.

e-Education

Alongside e-Health, there are also potential uses for this research within e-Education. For example, the use of dynamic constant-authentication would allow a system to keep monitoring a child during work. For example, a child's eyes could be monitored during reading to find out

if they look away for a large period which could intimate either boredom or misunderstanding, providing the teacher with a tool to improve the students work by providing help where it is most needed.

2.4.3 Research Forms

Along with the potential uses and users of this research, other factors need to be considered, such as the development of suitable understanding within the area being researched. While there is a basis of knowledge, the depth in which the research goes demands a thorough understanding of both the main area and all potential sub-areas. The following section will highlight this process.

When researching "Novel approaches to biometric security with an emphasis on liveness and coercion detection" there will be a lot of secondary research which is gathered from books, journals, conferences and other data stores and will provide an extensive background knowledge level that can then be used to expand in within the research proper. This literature review will identify the salient works within the different areas and will ask the following questions of the different topics e.g. biometric, liveness, coercion, scrutiny, taxonomy, etc.:

1. Are the topics clearly identified as important?
2. Are appropriate questions asked that link to this research?
3. Are the sources clear, specific and answerable?
4. Are they interconnected (if there is more than one)?
5. Are the sources substantively relevant (that is, interesting and worthwhile)?
6. Is the relevance to some theoretical concern, or some professional or practical concern, made clear?
7. Does it draw on relevant concepts from the relevant disciplinary literature?

The other main area of data collection will be a testing phase; this phase will involve the theoretical implementation of the taxonomy and algorithm and then a selection of tests and case studies to check the effectiveness of the research as mentioned in the grounded theory section.

2.5 Research Process

Throughout this section, there will be links to the objectives identified within chapter 1 and will confirm to the following format (Objective A.1.).

2.5.1 Background Gathering

The first stage of the research is to gather and disseminate current research in the areas of biometric security; this will comprise of developing a thorough understanding regarding the main factors surrounding the main areas of this research. The first stage will be to gather information pertaining to biometric systems. This will involve the understanding of the general biometric practice, the architectures involved, strengths and weakness of devices and techniques amongst other factors. As this is the initial stage, a lot of information will need to be identified and worked through to achieve the level of knowledge enabling the development of future areas (Objective A.2.). The next stage is to further the initial research and literature review to identify the other important aspect of the research, which primarily focuses on autonomic environments. This will allow an informed opinion to be developed that can identify if the integration of biometric techniques within autonomic environments is practical (Objective A.3.).

2.5.2 Characteristic Classification

After this knowledge has been gathered the dissemination of the factors will take place, which involves the identification of areas in which the different research areas could be incorporated. It will be at this point that the refinement of the overall research will occur, as certain factors will become more and more prevalent. The current expectation is that liveness detection and the integration within autonomic environments of biometric devices, in general, will be the most important areas, which once again links to (Objectives A.1. and A.3). Therefore, the focus of this stage will be to identify how these factors can be combined and what specific issues may occur when this combination is evident.

The main postulation is:

1. A taxonomy is a very important tool to develop for liveness detection and potential coercion detection as the current method of classification is not suitable due the lack of precision it provides.

2.5.3 Taxonomical Differences

Using the knowledge gained in the previous sections the taxonomical development will begin. The gathering and analysing of the current classification system will show the importance of a thorough and precise taxonomy especially if it is to be applied to liveness detection techniques. The previous section has will hopefully identify that the current classification system is not suitable and the taxonomy development is needed providing impact for developers and researchers (Objective B.3.).

The next phase will be to develop the requirements for this new taxonomy by identifying which factors are important, and which are not, which will be achieved by disseminating the different characteristics of liveness detection techniques. Throughout this process, the classifiers must be kept as general as possible as the intent is to use the taxonomy for other biometric techniques, maintaining a flexible and scalable taxonomy, (Objective B.1, 2.). When these classifiers are finalised, they will be mapped to a mathematical equation that will output a value that can be used to denote comparable security between two systems (Objective B.3.).

2.5.4 Coercion Detection Development

Initial research indicates that alongside liveness detection a second security feature is going to become important. Therefore, during the development of the taxonomy, analysis of potential integration for coercion detection technique will be validated. Coercion detection has been identified as the next area in biometrics by [17] and currently is a security flaw that can be bypassed with ease. Because of this, it is important to research further as it would be able to push the boundaries of knowledge in this area. This coercion detection research will be conducted by considering the current minor research, alongside the more fully researched precursors that identified the area as one not only woefully under-represented but also becoming more and more relevant, (Objective D.).

This new stage will identify the underlying structure of coercion detection, discuss how the technique could work by comparing it with both liveness and biometric security, architecture and process. This will then be analysed and a selection of applicable techniques will be developed which will potentially be original and novel, as well as some rooted more deeply within the biometric area overall, (Objective E.1.).

2.5.5 Taxonomical Suitability

The next stage will be to identify how easily coercion detection would be to implement into the developed taxonomy. This is the main reason the taxonomy is to be created as flexible as possible, so that if new and innovative areas are identified then they can easily be incorporated. The emphasis is to leave the taxonomy as general as possible from a head-on view, it will be possible to change the specific classifier meanings to adapt to the techniques being identified, (Objective B.1.) After this taxonomy, translation is complete the overall security value, as identified by the algorithm output, will once again be generated to check the overall compatibility to check for applicability, (Objective B.3.).

2.5.6 Taxonomy Testing

After the development of the taxonomy and the identification of additional relevant factors, such as coercion, the next stage will be to test validity, which will be achieved by inputting a selection of test data as identified by (Objective G.2.). The main purpose of this is to allow a variety of ways for a developer or researcher to utilise the data they enter. Subsequently, the testing will identify the overall ranking of different techniques, the different classifier totals, as well as specific characteristics of each classifier allowing a thorough understanding of the technique and to provide the user with a variety of useful information. During this testing stage both the validity of the technique and the identification of any problem areas/points of interest will be sought. This linked directly to (Objectives B.3.).

2.5.7 Algorithmic Development and Application

After the taxonomy has been developed the analysis of different entries into the taxonomy will be an important stage to consider. To compare these techniques an algorithm will be developed that takes the information within the taxonomy, along with scenario-specific data such as quantity of users and amount of access attempts, and generates a value that will denote the level of security that setup provides. By changing the taxonomy values for the detection technique, the algorithm also changes and subsequently the overall output will be different. This will be used to highlight what techniques work best together within unimodal and multi-modal systems as well as improving the understanding of fusion intra-technique. This algorithm will use smaller metrics that are development throughout the taxonomy to create the final overall result.

The algorithm development process will be hybrid using both brute force and divide and conquer algorithms. Brute force is being used initially as it is one of the most used

techniques, as the implementation of this algorithm will use a variety of different techniques and disciplines this brute force technique will be especially useful. Secondly, as the overall algorithmic complexity is not high, especially from a mathematical standpoint, the comparative simplicity of the brute force technique will be ideally suited to this development style [150]. The second algorithm being used will be the divide and conquer algorithms, and this is because the innate development will mainly consist of small algorithms that make up the overall main algorithm [75]. Again the innate development process is denoting the algorithm development process. This combination also, theoretically, negates some of the negative factors associated with the other algorithm type, for example, brute force does not use shortcuts and can be quite inefficient, whilst divide and conquers primary advantages is the efficiency it brings to the development process [118]. Whilst other techniques, such as the decrease and conquer and transform and conquer, could potentially be viable in some way, the focus on simplifying the problem to get a simpler result is not what is being sought in this research [11].

2.5.8 Algorithm Analysis

The final stage of this research will revolve around the final outcomes of the algorithm which can be tested to see if they work, therefore allowing a system to change autonomically depending on the information provided. This would also be coupled with an overall evaluation of the different aspects of the research, identifying the areas that need addressing within the taxonomy, algorithm and autonomic inclusion. This would include the evaluation of the taxonomy, (Objective G.3.), as well as other aspects such as the algorithm. These factors would identify if the overall purpose of the research has been identified, and if so how would it be improved, or what areas might cause for concern in the future, (Objective F.) and overall Aim. This would lead directly into the future work section.

Throughout this research, there will be numerous times that a factor deserves further research, but due to the scope of this research, it will not be practical to do so. However, within this final section, these areas will be identified and a brief identification of how they could impact the overall research conducted will be provided. This section, potentially, will show that whilst a substantial piece of research had been conducted there are still avenues to further understanding and, hopefully, that the research process has highlighted new and exciting areas.

2.6 Ethical Considerations

As this is theoretical research, and that there is no deployment with participants, there are minor ethical considerations to identify. There will be no testing of end users, no any data stored about users.

Subsequently, this research conforms to the Framework for Research Ethics [134]. However, due to the limited contact to external parties, most of the factors are not applicable in this case.

2.7 Research Risks

As with any form of research the associated risks must be considered, identified, and planned for Table A.2, identifies some of the main risks associated with this research. This risk assessment will enable any potential risks to be catalogued and good safeguards put into use so that they would have the least amount of impact as possible. This is especially important for the high score areas as a good preventative plan will minimise any potential disruption. Additionally, though the research process at regular intervals the risk table will be consulted and updated to represent any potential risks that might occur at a later date.

2.8 Conclusion

Throughout this chapter, multiple methodologies have been identified and the ones chosen have been justified within the scope of this research. The different stages of the research have been identified, including the background knowledge gain, taxonomy development, liveness, coercion and algorithm development and justification and a final evaluation and future work consideration. Throughout the subsequent chapters, the different characteristics of this research will be identified and along the way the novel and original contribution to knowledge will be highlighted to show that the overall process followed is producing research that applies to real-world scenarios.

The subsequent chapter will discuss the background literature that has been gathered and identified as important.

Chapter 3

Biometric Background

Blip *Blip* *Blip* End of Cheese Error
Blip *Blip* *Blip* Can Not Find Drive Z:
Blip *Blip* *Blip* Unknown Application Error
Blip *Blip* *Blip* Please Reboot Universe
Blip *Blip* *Blip* Year Of The Sloth *Blip* *Blip* *Blip*

Pratchett, 2004

The surge of biometric research, in recent years, has led to numerous innovations in the field e.g. sample collection from body odour, exhalation, ear print, etc. [104]. To gauge the salient factors, this section will consider the current state of biometric security. Using [67] definition: biometrics are "the automated use of physiological or behavioural characteristics to determine or verify identity". However as with many definitions further research has elaborated on this concept and [114] has added chemical characteristics.

This chapter will look into biometric security, how it works, what factors make it up and what areas need improvement. To begin the process of biometric security will be covered which will then lead to the different threats biometrics face. This will lead to liveness detection its techniques and flaws, finally culminating in the identification of coercion detection and the importance it has within biometric security.

3.1 Biometric Security

Biometric devices have enjoyed much greater levels of integration and research in recent years, partially due to the ubiquity of the technology, especially within smart device environments, as well as a general increase in acceptance of the technology. The major driving factor is the integration in smart devices, which range from fingerprint scanners in phones to iHomes

using a suite of biometric sensors [172]. Due to the mobility and ubiquity of these devices, the demand for dynamic and adaptive security is becoming apparent and it is here that biometrics can easily be integrated. Of course there are factors to consider such as: user acceptance, liveness detection [7] [166], coercion detection and security threat vectors.

Even though the use of these biometric devices has become more acceptable, according to [41], it still has yet to gain universal acceptance primarily due to the natural suspicion of this medium and concerns with privacy and security [181]. In 2005 only 15% of organisations within the U.S.A were using biometric devices, according to CSI/FBI Computer Crime Survey [65] and while this figure increased to 20.5% in 2010 it showed a decline of roughly 5% from 2009 [161]. This data shows that even though the potential benefits from biometric security are high, the acceptance and inclusion is not increasing as much as it potentially could which can, in part, be contributed to poor media representation and fears for privacy [28].

A major factor in the acceptance of biometric devices is the manner in which the chosen biometric reading, henceforth known as the sample, is collected and the form the collection takes. As the range of biometric techniques grows so does the following factors:

1. Is the sample gathering invasive
2. Is the process transparent
3. How long does it take to gather the sample
4. Is the sample secure
5. Can the sample be used for other things
6. Can the sample be removed if compromised

To answer these questions biometric devices have different characteristics that highlight their salient strengths and weaknesses. These categories are diverse and need to be considered for liveness detection, which deals with the spoof-ability of samples, as well as for coercion detection, which checks the corporeal nature of the sample. The following section will identify and discuss some of the most used and researched biometric techniques as identified by leading area specialists such as [81] [41]. Whilst there are numerous biometric devices to choose from, each having its own salient positive and negative features, this discussion will focus on the following four:

1. Fingerprint

2. Iris/Retina scanning
3. Facial Recognition
4. Vocal recognition.

These devices have different strengths and weaknesses and currently the main way to identify these is by using the ordinal measuring system as shown within Table 3.1, which is based on [79]'s work and Table 3.1 and Table 3.2 will form the core of testable techniques within the taxonomy.

Biometric	Universality	Uniqueness	Permanence	Collectabilty	Performance	Acceptability	Circumvention
Face	H	L	M	H	L	H	L
Fingerprint	M	M	H	M	H	M	H
Iris	H	H	H	M	H	L	H
Retina	H	L	M	L	H	L	H
Voice	M	H	L	M	L	H	L

Table 3.1 Ordinal classification of biometric devices [81] [130]

The same classification also applies to the host of other techniques that are available, and whilst the techniques identified earlier are the most prevalent within the industry: new techniques and adaptations of existing ones are constantly being identified and developed. Therefore, a selection of the novel techniques will also be identified so that the taxonomy can create a broader data-set. Table 3.2 shows some of the main non-common techniques [79] [175].

Biometric	Universality	Uniqueness	Permanence	Collectabilty	Performance	Acceptability	Circumvention
Keystroke Dynamics	L	M	L	M	L	M	M
Hand Vein	M	H	M	M	M	M	H
Signature	L	L	L	H	L	H	L
Gait	H	M	L	M	M	L	H
Ear	H	M	H	L	M	M	M

Table 3.2 Novel Biometric characteristics identified by [82] [176]

Unlike traditional passkey, biometric devices can use a variety of samples when authenticating. Therefore, one of the main factors to consider the method of sample acquisition and the form the sample will take. The first consideration is the form the sample takes as this will denote potential threats that can be more or less pertinent. These samples consist of biological and behavioural data, with biological samples being the norm and the more accurate style

[81]. These samples can consist of chemical, physiological, or medical based data. To identify what the salient factors of different samples have, an ordinal classification system provides seven different categories which grade a biometric technique. Whilst this allows a basic comparison the reliance on the ordinal measure is a problem, and will be discussed later in this chapter. These classifications provide information regarding the capabilities of the biometric technique. However, it is not suitable to apply to liveness and coercion techniques due to the innate difference with general biometric authentication and the lack of comparative analysis that a different measurement system would provide. To further apply these characteristics, especially to liveness and coercion detection, a more suitable measuring standard will have to be considered.

These classifications identify the most apt biometric to use in a situation, by having a simple and effective access to their salient features. The classification of biometrics is split into seven specific areas that endeavour to provide all of the relevant information regarding the overall level of security a technique has. Therefore, a brief description of the different biometric categories will be identified with a view to enable a thorough development of the overall taxonomy. It may be become apparent that some classifications are not suitable for inclusion with a different measuring system, and for liveness and coercion detection.

Considering the current measurement system, there are seven different factors each denoting a different factor of a biometric device. Firstly, the probability that a user has a specific biometric feature, which denotes the level of *universality*. For example, as Table 3.1 indicates, facial biometrics have a high universality as it can be safely assumed that a user will have a face and the chance that this is not the case is very limited. Alternatively, when considering signature biometrics the chance for a sample to be unavailable is much higher, unavailability can be caused by sample deviation, sample acquisition errors, etc., therefore making it a less universal technique. This is due to the innate transient nature of signatures in general and the potential issues that can effect signatures such as the lack of the relevant hand, potential neurological damage (stroke, dementia, Parkinson's disease, etc.) [35].

The heterogeneity of a biometric device is vital to enabling efficient use fusion architectures. Fusion is denoting when two or more devices or techniques are being incorporated together to improve the overall security. This also highlights how individual a sample is. The *Uniqueness* of a sample highlights how accurate it is when matched, as every technique will have a minimum level of comparison that shows if they are unique and if not how many people are likely to have the sample in a set number of users. For example: if a sample has a matching level of 1 in 100 it means that 1 in every 100 samples will be the same therefore representing a low range of security. If the technique instead is 1 in 1000000, then the overall security is much higher and therefore much more relevant.

The third measurement is *Permanence*, as all samples have degrees of transience that must be contended with especially when the original sample is derived from the human body and human behaviour. Permanence altering factors can occur due to a variety of factors such as accident to age degradation. The degree to which a sample is permanent and the relevant variance must be considered. Many factors, especially behavioural ones, can change dramatically over time, and can differ from day to day depending on some factors that are often beyond the control of both the user and the system. Examples include weather variations causing voice and signature problems due to excessive coldness: medical/impairment considerations due to either illness or age, such as cataracts within an iris scanning, circulatory problems within hand geometry, etc. [137] [168].

The fourth factor highlights the *Collectability* of samples. This identifies the complexity with gathering data, for example, a fingerprint scanner needs specific hardware to gather samples, therefore, it is a medium grade. However facial recognition only requires basic video and is a lot easier to gather, denoting a high collectability.

The fifth factor identifies the *Performance* of the specific technique. This considers different areas and the primary focus in this metric using False Accept Rate (FAR), False Reject Rate (FRR), Failure To Enrol (FTE), Failure To Acquire (FTA), etc. [171] [106]. Within these standards, the FAR is the more important [182] as it shows the acceptance of the false attempts upon the system and it is more important to keep the amount of false acceptance down than it is to keep the amount of false rejections down. Whilst there are some researchers that try and interject custom standards in these research areas [111] these factors are rarely followed and the standardisation process is highlighted as a positive factor [154]. This is one of the main problems with standardisation, if people create new measurements, then they add to the lack of analysis problem.

The sixth factor highlights the acceptability of a biometric technique and can dramatically impact the effectiveness of an installation. This characteristic can be influenced by some factors, from negative representation in the media: to ignorance of the technique and associated fear or mistrust.

The final characteristic denotes the ability to circumvent the biometric technique and the ability to create spoof samples. This is one of the most important characteristics and one of the main techniques to prevent sample spoofing in this way is to use liveness and coercion detection.

All of these characteristics link to biometric techniques as indicated within Table 3.1 and Table 3.2, and this shows that due to the range of biometric techniques, the correct choice of techniques depends solely on the system requirements. Therefore, a complete understanding regarding the techniques associated with the relevant devices is imperative. It

is because of this wide diversity that different liveness technique has become apparent and must be considered regarding their suitability within a security environment. One pertinent factor is the mutability of the sample, or its permanence, as different groupings of biometric samples can differ to a lesser or greater extent. For example behavioural characteristics are more concerned with gathering data that identifies normal behaviour for the user, such as keystroke dynamics, which can differ dramatically depending on mood, concentration, etc. [18] [62]. This data can also be used to highlight other factors surrounding the user, such as time taken to complete tasks, the speed of typing, etc. In the same area, medical biological characteristics can provide medical data that could be used in other areas such as health monitoring or product testing [140].

As the focus of this research is on biometric security the main tenets of authentication, authorisation and verification are key processes. [2] identified that biometrics can verify the user and also can discover the identity of the user [135] [41]. The verification process compares the provided sample with the user's stored template, then, depending on the accuracy of the sample, the user is provided or denied access. This uses a one-to-one process as the sample is being compared to a specific template. Identification uses a different technique as it compares the user's sample with a template database, and therefore will take much longer to complete due to the amount of factors being tested.

Therefore, the following sections will cover some salient points: firstly the biometric architecture will be covered, this is because to understand how certain factors such as fusion and security threats the process of biometric security must be highlighted. This will then be followed by the threats associated with this process and the attack vectors that are of chief concern. The third section will cover the process to prevent the most problematic threat vectors, the development of spoof data. This section will highlight the ways in which liveness minimises the effect of spoof data but is susceptible to coercion based spoof data. Subsequently, the final section will discuss the importance of coercion detection and what issues are currently most prominent in the area.

3.2 Biometric Architecture

There are five main modules within the biometric authentication: sensor, feature extractor, template database, matcher and decision module [103]. Each process provides specific information that is used in the subsequent processes as shown in Figure 3.1.

The sensor is the main interface for the device and is the primary form of contact between the users and the system. The function of the sensor is to scan the user's relevant biometric data which may differ depending on the devices used, and sample requested e.g. fingerprint,

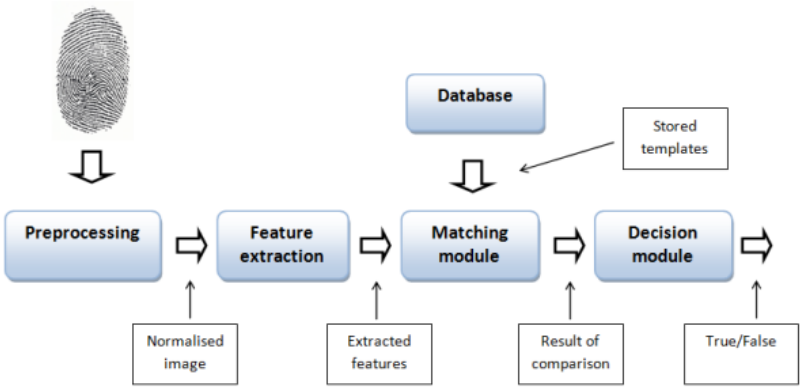


Fig. 3.1 Biometric Architecture

facial scan, etc. [103]. When gathering these samples, user acceptance is a key factor as the process may be invasive and there may be fears for the data security. The technique used should address these concerns. However general improvement of user opinion and knowledge is a key factor. One method of doing this is to make sure that the interface is as usable as possible and conforms with all accessibility and usability requirements that any other interface would follow. Another factor to consider when gathering the sample is its universality. Therefore adequate provision must be provided to facilitate any users that have specific needs such as those that have any impairments that may inhibit, prevent, or simply make the authentication process more difficult. This is a very important area as each technique of biometric authentication i.e. fingerprints, vocal recognition, facial recognition, etc. could cause specific impairments to be problematic during the sample gathering period [41]. The sample is then forwarded to the next process which is the feature extractor.

The next process takes the captured data and gathers the salient features of the sample which are to be compared. The feature extractor creates a map of its unique features, for example, the features could be the distance between the ridges and valleys on a fingerprint or the distance between specific minutia. These features are what will be used in later phases to authenticate correctly or verify identity. Depending on the device being used there may be an additional stage which checks if a provided sample is of sufficient quality for the feature extractor to gather sufficiently usable data [41] [80]. This stage can suffer from permanence altering factors. For example, vocal recognition can dramatically change depending on the quality of the device, environmental factors, current medical factors, etc. Therefore, the addition of a quality process can improve the overall process dramatically. After this feature extraction has happened the next stage is to compare the map with the stored sample, known as the template.

The template database holds all fo the templates for the authorised users. If the process initialised is an enrolment then this is where the final template will be stored for future comparisons. When authenticating a user the extracting features are compared to the pre-provided template. The template database is a very important area for both the authentication process and from a security point of view as it stores all of the official template: Firstly this is an area that is ripe for a malicious attacker to target, this is primary because all of the templates are stored within this database allowing a malicious user the ability to recreate templates exactly, edit templates to fit with other users, or delete templates therefore preventing users accessing their login credentials. Another issue is the more users denotes more storage needed for the templates, and subsequently, a more likelihood for administration attacks to occur due to administrative negligence [103] [41]. When the data from both the feature extractor and template database has been gathered the next stage is to match the data sets.

The matcher highlights the difference and similarities between the two samples provided. This is done by taking the two biometric feature sets; one being the stored template (Xt) and the other being the gathered sample (Xq) this, in turn, outputs a score (s) that denotes the similarity between the two templates. This score, in turn, is sent to the final module for a decision to be made [41].

The final phase of this process is to take the score (s) that has been generated and make a decision which allows or denies the user access to the system depending on the (s) and its mathematical similarity to a predefined acceptable limit. This limit can, and does, differ intra-device as the salient features within different biometric techniques can be quite different. If the score is mathematically similar enough to the template, then the user will be authorised into the system, however if there is insufficient similarity the user will be denied access[80] [41].

While these processes are quite simple individually the amount of threats that can occur is extensive and this becomes increasingly more problematic when considering other factors such as multi-modal architectures and the differing fusion locations. Therefore to understand the threats that are occurring the following section will highlight some of the most impactful factors.

3.3 Biometric Security and Privacy

Biometric security is often modelled using Ishikawa diagrams (also known as fish-bone or herringbone models), which identifies the source causes for system vulnerabilities. While these diagrams provides a very clear and comprehensive outline of system vulnerabilities

it does suffer from a reduction in clarity as diagram data is increased [71]. Using these diagrams: two forms of failure are highlighted which cover most of a biometric system security threats, these are intrinsic failures and adversary attacks. The following section will cover these threats in more detail.

3.3.1 Intrinsic failure

Intrinsic failure is caused when the system incorrectly completes a process such as authentication, enrolment, etc. The two main forms are false reject and the false accept states [113]. False rejects occur when a correctly enroled user is denied access to the biometric system, normally because when the matching process has occurred the score is not within an acceptable range for the matcher to accept. This can be caused by intra-class problems such as sample divergence or noise. This can be caused from internal problems such as corrupted templates but can also be an external influence such as dirt on scanners or residual scans from previous users etc.; this is especially problematic for finger/palm scans [130]. False accepts are normally a product of poor data capture, the query does not have enough relevant data to be used within the matching sequences, this lack of uniqueness can cause issues in authentication throughout the system as a lack of uniqueness can cause data to be misidentified as other biometric templates [136]. False accept rate highlight the opposite problem occurring. A spoof sample is incorrectly authenticated into the system, therefore, allowing a nefarious user access. This FA is the more damaging of the two factors, as whilst the denial of service caused by FR will cause efficiency problems, there is no actual security breach. Unlike FA, which allows a user into the system that should not be there, circumventing all of the security in place.

Depending on the quality of the biometric scanner other issues can occur, such as FTE (Failure to enrol) and FTA (Failure to acquire) errors [115]. FTE problems occur when the user is providing the initial sample for use within the template database, whilst FTA occurs when the system cannot take the sample provided. This can be when enroling or authenticating; it is an all-encompassing metric. This might be due to limitations with the technology such as palm scanner cannot accept user with smaller or larger hands therefore creating either an FTA or FTE state. This problem becomes exponentially larger when there are more people using the system. A small department with ten users may have few problems as it is easy for the system to be developed specifically with the users in mind. If the target audience is a university campus containing five thousand plus users, then the chance for FTE and FTA to occur increases exponentially. Similar issues can occur if the users have any kind of disability that may prevent the user from authenticating or that causes errors to occur.

Examples can include visual impairments such as blindness, Anophthalmia, cataracts, etc. [137].

3.3.2 Adversary Attacks

Whilst intrinsic attacks deal with threats that emanate from within the system and therefore do not have an nefarious external instigator; adversary attacks are just the opposite. These attacks occur when third parties attempt to infiltrate the system. Adversary attacks are dependent on the discovery and exploration of loopholes or errors in the systems security. These vectors are often found in the intra-process areas between different architectural modules such as the feature extractor and template database.

There are a variety of vectors that can be affected. Firstly, as all systems need maintenance and administration to function correctly some error, mishandling of procedure or simple fault in the integrity of the data due to incorrect administration can cause the relevant loophole to occur for the attacker. These threats can be caused by a plethora of different reasons, for example, incorrect enrolment: e.g. the systems administrator provides greater system access than is necessary for a new enrollee. There are also issues that are caused by coercion that may be centred on either a systems administrator or a legitimately enroled user being forced to allow a non-authorised user access. This is an area that has very little discussion but massive impact, as liveness detection is considered within threat vectors but the effects of coercion are not [131].

This culminates with biometric threat vectors, which endeavour to identify all possible areas that an attack could occur and which will be most devastating. [131] postulated that there were eight points of attack in biometric system and whilst these points are quite general they do encapsulate the main threats that a system would be susceptible to as shown in Figure 3.2. This did not identify device specific threats instead focusing on a generalisation [131]. Whilst this initial model was valid, it became obvious that it did not go into sufficient detail on the potential vectors and therefore [175] [45] expanding on these ideas.

Figure 3.2 identifies the eight threat vectors that [130] described. The initial attack covered the attempted enrolment using a fake biometric sample; this could include fingerprint, facial recognition, signature theft etc. such as a gummy finger creation, video/image spoofing, vocal recording as discussed, respectively, by [171] [105] [162]. Solutions such as improved quality of scanners and integration of liveness detection were deployed to combat this threat, and have been successful to a point. However there are still worrying amounts of biometric sensor being shipped without, or with sub-par, liveness standards [120] [166].

The next vector covers the re-use of stored digital biometric samples. A nefarious attacker would record the signals between the scanner and the matcher and the provide this

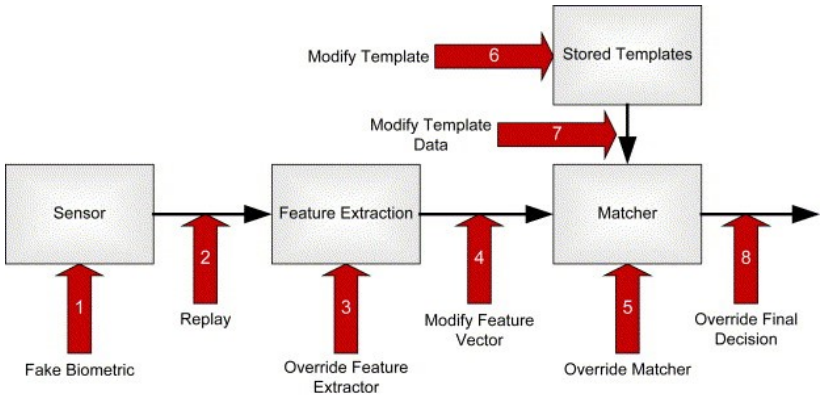


Fig. 3.2 Biometric threat vectors according to [131]

data, instead of an actual sample, to the matcher module bypassing the scanner entirely [7]. Alternatively, the data could be sent to the feature extraction process, therefore bypassing the initial sample collection [130].

The next vector overrides the feature extractor data replacing it with ones provided by the attacker. This can be difficult to do as the feature extraction method and matcher are often very interlinked and therefore separating them enough to attack in this manner can be problematic. It can be done much more easily if there is some form of remote matching system as this would allow the separation and time the attacker needs [82].

The matcher is an ideal place to attack, as anyone with control over this section can provide a correct match to any sample provided Therefore bypassing all other security aspects, this is partially why data is often encrypted by using salting or transform schemas within the template database so that it can minimise any matcher security problems [7].

As with any data stores, it creates a very tempting target as identified by [130], highlighting that any attacks on the template database would be very problematic. This is because there is a likelihood that the template database is remotely stored, due to redundancy, backup, etc., therefore creating a more easily accessed area for an attacker to target.

Template attacks can be some of the most devastating to a system, due to the sensitive data held within this area. There are techniques to minimise any potential damage, but there is a need for a set of standards that can provide the best security for a template system. [115] hypothesises that there are four distinct areas enabling as close to ideal as possible protection scheme. These four areas are, diversity, revocability, security and performance [126].

Diversity identifies that there must be a secure template that does not allow cross database matching as this could interfere with the privacy of the user, and impact the diversity of the sample and template. Revocability controls how easy it is to revoke the template, often a problem area with biometrics due to the limited entropy of sample. This revocability decrees

that should a problem occur to a user's data, therefore compromising it; then it should be easy to cancel or revoke the data and resupply the user with another authentication key using the same piece of enroled data, i.e. using a different selection of finger minutiae as the matching criteria. This template entropy is a key factor within the biometric security and valid understanding of the different techniques that allow one sample to be split into multiple potential templates. Following this, the security of the template is one of the more obvious aspects that needs consideration. Making it is difficult as possible for a nefarious attacker to gather actual template data which could be used for a host of nefarious purposes such as: changing the stolen sample; make spoof samples; deleting valid templates etc. Finally, the performance of the actual template matching and schema must be taken into account, whatever protection schema is selected should not adversely affect the general performance of the biometric device specifically when considering the FAR (False Accept Rate) and FRR (False Reject Rate) [82].

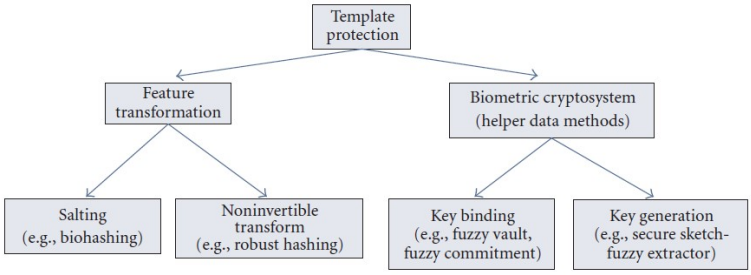


Fig. 3.3 Template protection categories [115]

To achieve these template protection goals, there are numerous protection schemas which provide different characteristics and advantages. Some of the more well-known schema are identified within Figure 3.3. After passing template protection, the authentication process must provide the template to the matcher, therefore, allowing the module to authenticate the user. If this vector is compromised the template can potentially be substituted or changed, however it can be protected somewhat by using salting and transformation protection schemas due to the added encryption/transformation robustness and comparability they provide.

The final vector [130] discussed has the potential to be the most devastating attack but is also the most difficult to achieve. If an attacker can influence the final decision of the matcher, possibly by intercepting the message sent from the matcher to the decision module, then it would render all other security techniques irrelevant as the system would be working correctly but the final decision would not represent the official data.

These initial threat vector whilst valid, were over-simplified and subsequently over the years more advanced and thorough vectors have been identified, partially due to technological

improvements, and partially due to the maturing of the subject area. [81] has identified five different subsections of a general biometric systems each with a verity of threat vectors as shown in Figure 3.4.

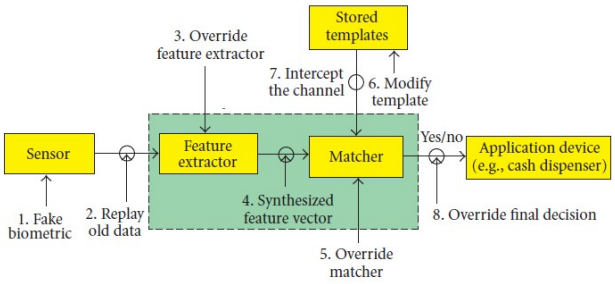


Fig. 3.4 Expanded threat vectors based [130] [114]'s work

These additional vectors encapsulated many of the originals shown in [130]'s work and are shown in Figure 3.4. These have been further researched by [114] who identified four vector categories, as shown in Figure 3.4 and [21] who further extended this work by creating a framework that was based around [176]'s five sub-sector model. This made a framework with twenty attack points and twenty-two vulnerabilities as shown in Figure 3.5.

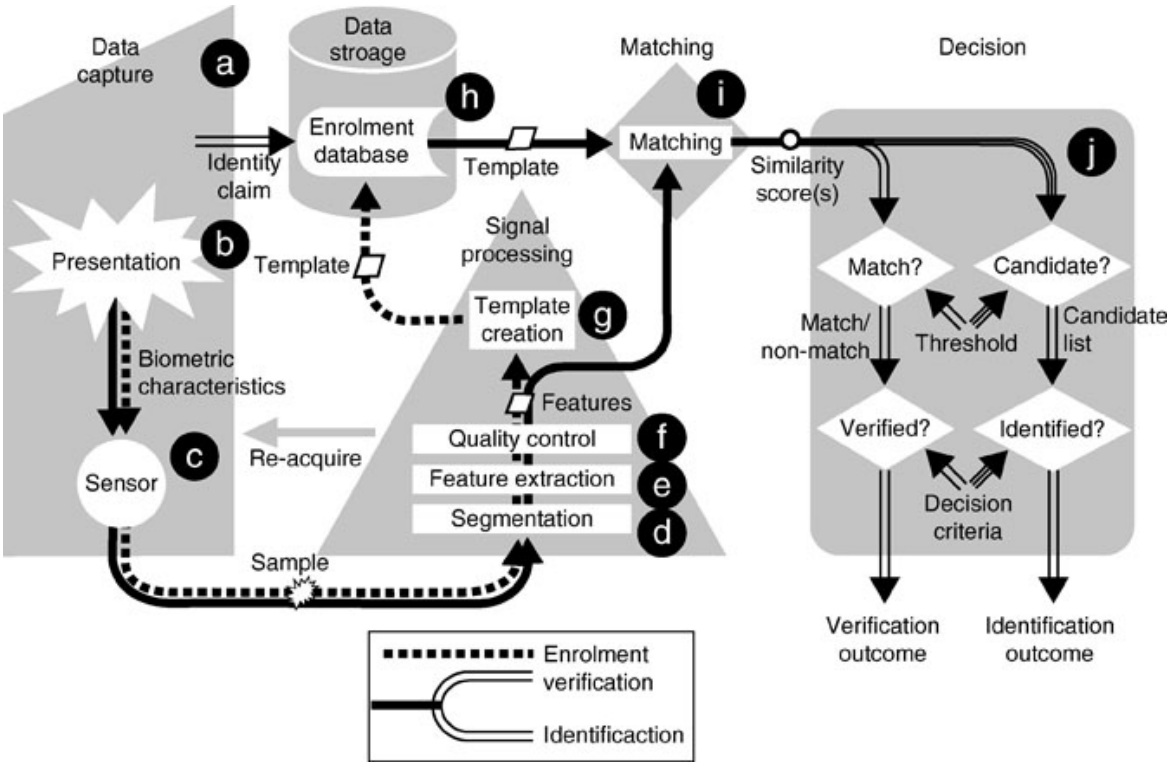


Fig. 3.5 Bartlow and Cukic's threat expanded framework [20]

[114] identified four areas in which attacks might occur and these are; user interface attacks, inter-modal attack, module attack and template database attacks as shown in Figure 3.4. User interface attacks involve the use of a fake or spoof biometric characteristic which could consist of a variety of techniques such as gummy fingers, HD imagery and so on. The second vector addresses the potential security issues within the system that occurs due to the different modules/entity communication. Potential attacks could occur from two main threat techniques: remote interception of signals or from a local jammer/interceptor. These attacks are conducted by a nefarious user who substitutes their own signal for example, changing a sample or changing the location of a correct template and replacing it with a spoofed template. This Trojan horse like attacks can be countered by the use of secure software practices and the inclusion of specialised hardware/software which will improve the chance that the algorithmic integrity of the process is maintained. The fourth potential attack vector, template database attack, has the potential to be one of the most devastating as the template database is where all of the enrolment samples are kept ready for matching. There are some issues that can arise from this kind of security breach: it allows the attacker to substitute a spoof biometric template with one that contains the attackers, or their representatives, sample. Subsequently, this gives the attacker the ability to access legitimately the system (from the matchers viewpoint) or the ability to create an official template, therefore, creating a legitimate way into the system that does not really on the initial attack, therefore becoming harder to both trace and solve. Table 3.3 indicates some basic biometric techniques and the data residue that is available, therefore facilitating the ease of developing biometric spoofs.

Device	Data used for enrolment	Potential data residue.	Ease of gathering
Fingerprint Scanner	Fingerprints	Residual prints on the scanner	Easy
		Prints from other surfaces e.g. glass, plastic	
		Discarded products	
Facial Recognition	Facial images	Camera Phones	Easy - Hard
		Camera	
		CCTV	
Vocal Recognition	Voice patterns	Sound recorder	Medium - Hard
		Mobile Phone	

Table 3.3 Residual biometric sample data collection

There are a number of important details within a biometric security implementation, such as how powerful a device must be to gather the best sample, what samples are most applicable to the situation, the susceptibility to noise, etc. As with many newer technologies such as biometrics and ubiquitous computing, there is a noted lack of coverage and standardisation.

Therefore, the development of a standardised process would improve the overall ease in which this security and installation factors could be identified and implemented. These standards would have to become significantly more prevalent if the areas are to be integrated within modern computing environments, especially as the emphasis on adaptation and dynamic technologies is becoming ubiquitous within the entire computing environment [41], however, this would partially rely on the willingness of manufacturers, such as LG and Sony [61], as well as researchers.

The main technique in combating these security concerns is to develop and implement liveness detection techniques, which are often seen as the main technique of sample validity and as [105] postulated liveness detection is possibly the most important security aspect, which was later backed up by [170] who proved that a device using liveness detection achieved higher success rates including false authentication drops from 90% to 10%. Liveness detection offers a number of advantages such as a way to minimise the impact of unauthorised data gathering, a common issue within biometrics. As Table 3.3 indicates some biometric techniques are prone to sample divestiture, the process of a security sample being gather-able from multiple places, for example a discarded coffee cup that is thrown away by a user. A nefarious party with this sample, and the knowledge of how to develop gummy/artificial fingers, as indicated by [171] [105], has the capability to develop a valid spoof sample. Liveness detection endeavours to reduce the effectiveness of these effects by requesting additional information about the living state of a sample. A further consideration is an effect that superfluous information can have on the accuracy and efficiency of the biometric [110]. This can be very important at a social level as certain biometric capture styles may cause undue levels of anxiety, stress or fear in a user due to any number of reasons such as ignorance of the process, fear of the technology and user coercion. A further consideration is that the samples are mutable and certain things like humidity, brightness etc. can affect the sample quality [41]. These problems are generally combined under one heading, noise, with some sub categories such environmental or medical etc. Finally, algorithms need to be used that will only gather the relevant data from a sample and discard the superfluous detritus (noise) such as described by [110] regarding iris recognition and the method in which to ignore noise such as eyelashes and eyebrows, facial occlusion, which is also backed up by [41].

Biometric devices can utilise many different techniques to achieve these features such as minimal distance, probabilistic methods and neural networks [41] [94]. [41] goes on to say that the neural networks are one of the most efficient as they learn certain characteristics about an authorised user so that it is quicker to match the templates and accept or deny that authentication request. This would be very useful and, if reliable, could improve the

security and efficiency of certain devices. Regardless there are some issues with neural networking as shown by [41] and [94]. Firstly the configuration of the neural networks is very complex and the computational requirements for large scale networks are very intense. This would make small-scale installations less viable and would require other concessions such as additional storage and specific training which would also be prohibitive depending on the size of organisation and userbase [41] [94].

Possibly the biggest risk for biometric systems is false physical and digital biometrics [135] of which there are numerous potential applications such as: using samples from a cadaver; coercing an unwilling living subject; lifting fingerprints from cups; and so on. Most of these methods take little technical knowledge or skill and can fool most simple biometric devices as has been proved on the television programme Mythbusters [1].

To combat this major problem, liveness detection was created. Liveness detection consists of checking for a signal that denotes the sample is from a living user and can be applied to all biometric techniques: heat; moisture; blinking; etc. [170]. This technique has improved security but is still an issue, as some of the signals can easily be fabricated such as heat and moisture, others, such as blinking whilst using a digital version of a retina scan can be more difficult to spoof but can be circumvented. Therefore, after the discussion of threat vectors an in-depth identification of liveness detection will be conducted.

Finally [65] discusses threat vulnerabilities, which elaborates on the plethora of generic system vulnerabilities, including OS, Storage, and Soft/Hard/Firm/Middleware issues. This is where more issues arise as biometric devices need additional security from physical attacks in addition to the normal software security parameters [135]. Normally the physical security of a system is comparatively simplistic, e.g. to lock the server in a password protected room. However a biometric device needs to be on public display for the users to access, subsequently becoming immediately vulnerable to physical based attacks and spoofing.

When developing these security installations, considerations must be given to the physical locations and to the variety of biometric techniques. There are many challenges and solutions that have already become well known within traditional security systems, as identified by [135] [156]. After the consideration of the physical implications the next defence consideration is liveness detection, which has been proved by the [170] and validated by [135] to improve the security of systems by up to 80% as it makes the authenticating user provides additional biometric information which can include perspiration, pulse, thermal and lip synchronisation. All these techniques are available to improve the security and provide answers to issues such as [106] “gummy ”fingers problem. Finally, there is multi-factor authentication, which involves the use of multiple forms of authentication such as tokens, pins, and smart cards, these methods deter spoofing attacks, as there are multiple forms of

authentication that need to be checked and there is not a single point of attack provide, as is often the case within biometrics [135].

Throughout this research it has become obvious that liveness detection is the key factor within biometric security, as it combats the main problem within biometric security, spoof samples. Therefore, the next section will discuss the salient factors within liveness detection.

3.4 Liveness Detection

The purpose of liveness detection is to verify and confirm that a sample being presented is from a correctly authenticated and living user [170]. The ideal is that liveness detection removes the problem of a user's sample being used by a third party whilst they are not there. This can be achieved by using spoof biometric samples gathered from their legitimate data sources such as lifting fingerprints, recorded voice recognition, or facial images for facial recognition [166] [3] [52].

This is necessary as it has been demonstrated that, whilst biometric security is very efficient and effective, there is a specific need to include liveness detection as spoof data is neither hard to gather nor hard to fabricate [82]. By including liveness detection, the risk of security breach due to spoofing attacks is dramatically reduced [166].

There are many methods and examples of liveness detection currently in operation. These include pulse, blood pressure or ECG [26]. Table 3.4 identifies some of the more well-known and acknowledged liveness detection techniques within biometrics along with the biometric it is applied to along with its advantages and disadvantages.

[ht]

Liveness Detection Technique	Advantages	Disadvantages	Device
Temperature	Easy to check. Easy to implement	Not secure Easy to spoof Lots of research on the topic [166])	Fingerprint Hand print Ear-print
Optical Properties	Very secure	Main forms of gummy spoofs have similar optical properties to skin Therefore minimising the use. [166]	Fingerprint
Spectrographic properties	Can detect a number of difference such as glass eyes, dead tissue	Expensive to implement	Hand print Ear-print
Pulse	Widely known technique Easy to gather.	Easy to spoof. Can differ dramatically due to host of factors – health/fitness of user, medical condition, recent activity etc.	Fingerprint Hand print Earprint
Pulse Oximetry	Medical application	Can be fooled by user providing their own sample covered with spoof i.e. fingerprint. Difficult to gather	Fingerprint Hand print Ear-print Vein Scan
Relative dielectric permittivity	Potentially very accurate Easy to implement (device dependent)	To get an accurate reading the admissible range would include that range of spoof samples	Fingerprint Handprint Vein Scan

Infra-red/Ultraviolet/ Thermal Scanners	Can be very effective at deterring spoof samples. Can show temperature, vein layout etc.	Very expensive They have to be installed and therefore ad-hoc authentication it not available.	All camera based
3d Head Movement	Prevents most basic forms of spoof (i.e. pictures of the user)	Can still be spoofed using video as a medium? Certain techniques are very prone to noise e.g. depth information vector [40] [121]	Facial
Micro Movements	Intuitive Comparison between the muscle movements when speaking occurs	Sometimes relies on additional hardware [121] Biometric dependent	Facial Blinking
Text Prompting	Integrated in the general form of vocal biometric styles. Very cheap to implement	Vulnerable to text to speech software	Vocal Facial
Hippus deviation	Little is known about it. Good at removing spoof attacks	Can be adversely effected during the ageing process [166]	Iris

Table 3.4 Liveness detection techniques identified by [166]

Whilst most forms of biometric device require liveness detection techniques to be added on, there are some biometrics that have inbuilt liveness detection due to the characteristic that is being looked for. An example is electrocardiograms. This is because to gather ECG data; then the user must have a living heart [26]. This liveness detection also allows the collection of certain medical data, which could then be used for a variety of other factors. For example monitoring the heart rate of a user to detect any abnormalities. However one factor must always be remembered, that whilst liveness detection brings a host of advantages to a system, it also brings a number of additional considerations as the system as is always the case when including an additional layer of complexity to a system. Issues that could affect liveness detection sub-systems include performance, circumvention, user acceptability etc. [136] [166] [7].

Liveness detection is normally processed during either the acquisition or processing stage of biometric authentication as these are the most relevant areas when dealing with a user's biometric input. This process is normally achieved by implementing the liveness detection in one of the following three methods [3][3].

Firstly, the inclusion of additional hardware which in turn can use the relevant data captured by the biometric device or, depending on the biometric in question, allowing the use of an already present liveness feature set that some biometric devices possess [166]. The primary disadvantage is that there is an acquisition of new hardware which introduces the problems associated with any new piece of hardware such as cost, size, heterogeneity and suitability issues, etc. One primary advantage would be the ability to add complex hardware, therefore, making the overall system more robust and allowing an easier integration of additional features [3]. Using this aspect within the proposed architecture provides a number of advantages. These range from more resources to use within the liveness module, the ability to incorporate additional hardware features that can capture other data, such as medical data, and the ability to incorporate a multi-modal liveness detection systems. However, as mentioned the primary disadvantages of this is that there is a spike in additional cost for the system, as well as maintenance costs and requirements. This method is especially pertinent when there are specific biometric requirements that may need additional hardware features such as, infra-red/ultraviolet camera inclusion, ECG testing, etc.

Secondly, using information that has already been captured by the original biometric device. This method is very cost effective as it does not require additional hardware [3]. The most important issue here is the complexity of the system needed to gather this liveness sample. This can be prohibitive to include therefore making additional hardware seem a more suitable option [166]]. Obviously, this method is also reliant on the applicability of the devices and the ability to gather the relevant information for the biometrics. Again

ECG would be valid as the authentication necessitates the liveness of a user, whilst a basic camera for facial recognition, may not have either the hardware complexity or algorithmic compatibility to provide the data needed for liveness detection.

Thirdly, by using inherent liveness features in the biometric sample, but which has not automatically been gathered when the sample has been provided. However, as this would require additional data collection, there are a number of problems when considering the heterogeneity, efficiency and user acceptance [166] [3]. There are a number of computational considerations such as what resources the system will need to collect this additional data, storage mediums and privacy concerns, all of which must be adequately addressed. This method is often viewed as a reactive concept as it shows the lack of forethought when gathering the original sample. Therefore, using a method that gathers data from the original biometric sample is a more elegant procedure, less invasive, and less resource dependent.

As well as different methods of gathering the liveness data, there are also three types of samples that can be gathered [181]. Firstly, there are intrinsic properties of a living body which can vary depending on the type of sample gathered. For example, there could be physical characteristics such as density and elasticity, electrical permutations such as the resistance and capacitance and spectral characteristics such as colour and opacity [166] [129]. Other examples include the deformation of the skin when pressed upon a scanner [37]), the relevant oxygenation levels of blood as seen through an ultraviolet scanner amongst others [121].

Secondly, there are the involuntary signals of a living body which can provide a lot of liveness detections data. Examples include perspiration, blood flow and ECG signals. Due to their inherent high level of security and built-in liveness detection these are often very highly thought off as security options [10]. As these samples are primarily medical in nature, it would be a small integrations to also allow involuntary liveness detection signals to be used for medical purposes as well, such as wellness detection potentially with integrations into smart home systems [166] [129]. However this medical link can also cause problems for user acceptance as the users may not wish to provide their medical data to the system.

The last area is the way the body reacts to external stimuli, unlike the previous techniques which gather data from internal effects on the body, this technique samples gathered from external stimuli invoking a response from the body. This is normally one of the most regularly used forms of liveness detection and is based on a challenge-response paradigm. This necessitates a cooperation between system and users. Examples include changing facial features, typing, changing of fingers for fingerprints etc.[166] [129].

All of these liveness detection techniques are designed to achieve a single goal; to prove the living status of an authenticating user. They represent a number of options relating

to implementation and suitability. However relevance may differ depending on specific scenarios [86]. A system with a high number of users may benefit more from a hardware based solution because of the robustness and quantity of users, whereas a small user base may not require the hardware solution [166].

As has been identified throughout this section liveness detection is a vital component of any biometric security system, but there are often a number of perceived issues with these instigations. Hardware is often difficult to manage and maintain, the collection process may not be robust enough to gather the relevant data at the point of enrolment, or there may need to be an additional level of authentication that must occur to gather this data, Therefore creating an invasive and visible system process, nether if which are ideal [166] [129]. These areas can produce a highly effective synergy between the biometric systems and liveness detection considerations that may be relevant [166] [86]. After a sample has been provided the first stage is to check if it is from a living user when this has been successful are there any other concerns to address? [166] identified that coercion is the next logical step to consider. However there has been little detail covered currently. In the following section, coercion detection will be discussed to highlight any factors that need considering.

This section has highlighted the need for a coherent and efficient method of comparison between liveness techniques. This is apparent due to the sheer variety of research and the variance of terminology and metrics within each individual research project. With the quantity of threat vectors and areas of spoofing, prevention is the key to improving the technology. Until there is an efficient method of categorisation and comparison research, there will be research that covers the same areas and uses individual measuring techniques. As [116] identifies the method of measurement within biometric system are the key techniques of comparison. Therefore, the main need is to create a system that will enable the same metrics to be used, allowing a greater comparison to be drawn.

3.5 Coercion Detection

Liveness detection was begotten from necessity after implementations of biometric environments were found to have a high degree of susceptibility to spoof data. This vulnerability demanded a robust defence and therefore liveness detection was developed [166] [81] with the intent to minimise spoof data effectiveness. Subsequent research has developed a variety of liveness techniques however in the manner of computing systems everywhere the threats to biometric security have adapted to take into account these new challenges.

Traditional key and pin based security systems have always had a number of threats one of the most noticeable and dangerous is phishing, the act of trying to gain secure information

by tricking users into believing the attacker has a legitimate right to ask for the information [156]. Whilst this problem has been around for many years people are still somewhat unaware of its importance and it has taken robust marketing on the behalf of companies to reduce the potential impact. These techniques normally revolve around a disclaimer claiming that 'A member of this bank will never ask for your passcode'. The intent here is to improve the knowledge of the user, however, even when passwords and codes are captured in this manner there are a number of factors that can be utilised to minimise any problems, such as password entropy [102]. When a passcode is captured it can be cancelled very easily, as the cancel-ability of any passcode is very high, alongside the ease of redistribution can be issued almost immediately. This, coupled with the potential for high degrees of password entropy, regardless of the actually implemented level of entropy [169] [139] employed by most users, are one of the prime reasons that passcode/words have had such a lasting presence within a security [156].

Biometric security does not conform to this structure and the lack of cancel-ability is a major disadvantage, as there are often a limited number of samples a user can provide due to a limited amount of fingers for example Therefore, there are a number of factors to improve the standard of biometric security, such as liveness detection and now coercion detection. Whilst liveness detection attempts to verify the corporeal nature of a sample, minimising the effects of spoof data, such as HD imagery, gummy fingers, and HD video, there has been minimal research and development conducted into coercion. In this case indicating that a user is, whilst living and therefore passing liveness detection techniques, authenticating under their own volition. Whilst this is a security threat itself, there are also features that, when used within biometrics, highlight the potential necessity for coercion detection. If a nefarious user attempts to gain access to a password based system they have to steal the password, for example employing phishing techniques. When this situation occurs the user is impacted, however, there is a degree of abstraction from the effect of the theft. However if the sample being used is attached to the user, then a host of other factors become paramount [143].

"To persuade (an unwilling person) to do something by using force or threats"is how [50] describes coercion. When considered in context, this means to force an unwilling user to utilise their biometric sample to gain access to a biometric system. Now that liveness detection has become a standard inclusion, this is the next stage of biometric security and must be considered as there are a number of important implications.

Firstly the act of coercion is abhorrent and all measures should be taken to prevent coercion being a viable tool, both from a moral and technical standpoint. Due to the nature of biometrics any nefarious entities wishing to access the system using someone else's sample can resort to coercion, therefore presenting an undesirable factor of the technology to the fore.

However, as some of the main forms of coercion detection utilise stress related physiological data there is potential to detect authorised users that are being coerced.

Due to the nature of coercion the focus must be on preventing any harm coming to the users, therefore making the comparative subtlety of a coercion detection technique highly important. Regardless the is little current research in this area therefore representing a challenge to develop novel techniques that area applicable. This is something that is being focused on within this research and a selection of both current and novel techniques that can be applied to coercion detection will be developed, alongside a critical evaluation of the primary issues that will affect this area [125].

Coercion detection has yet to be explored thoroughly as it has not been vital in previous incarnations of biometric systems due to two reasons:

1. Firstly, whilst coercion detection is important, the optimal and smooth running of the biometric systems coupled with liveness detection characteristics have been more so [114]. Therefore the focus has been on the development of these areas and improving them to a standard that is high enough for the research community to move on. Liveness detection has multiple implementations within public environments such the biometric cashes catering system installed at Oasis Academy [5] and Planet Cash, a biometric ATM based in Poland as discussed in Hitachi's white paper [74]. There are very few implementations or discussions on coercion detection and therefore due to the increasing ubiquity of liveness detection, more emphasis can be placed in the development of coercion detection standards and characteristics [143].

2. Secondly, most biometric systems have been implemented in secure environments such as airports, border checkpoints, etc. These areas already have additional security implications that are combined with biometrics to provide a more robust multi-disciplinary security environment [17]. Nevertheless, as increased integration of biometric devices and systems continues to invade personal computing, such as smart device integrations, smart homes and to build etc. and provides services throughout these environments, more consideration must be given to coercion. Along with the effects, it may have within the range of environments, it may need to be included in, for example, is it more likely to be a victim of coercion within a home or business environment [68].

These factors have shown that there is a need for coercion research and as one of the goals of this research is to develop a taxonomy that is flexible enough for future research integration it is a perfect environment to discuss coercion detection in detail. The primary focus is to identify what factors make the area as one major disadvantage is the lack of research and reasoning that has been conducted. There are very few works of research within

coercion detection and because of this it is very difficult to identify the salient features. [17] discusses some basic techniques of coercion detection in their project titled 'State of the art of mobile biometric, liveness and non-coercion detection', unfortunately, the depth of discussion is limited to four basic environments in which coercion detection can occur [68].

As this area is predominantly under-researched the development and categorisation of different coercion detection techniques is of paramount importance, not only to further the understanding within this area, but to provide actually applicable techniques that can be utilised within a system environment.

3.5.1 Development Background

Coercion detection is a very promising area that can have a profound impact on the degree of security within a variety of systems. As these techniques have progressed from liveness detection, a lot of the same considerations must be identified and contented with. The focus of this section is to identify the main concepts, and areas to consider within coercion detection. This will involve the development of techniques that in turn will be categorised using the taxonomy development covered later in this research. Therefore the first stage will include an identification of underlying concepts, and their relevance, and whilst this will not be an exhaustive list, due to the huge potential of options available, and the limitations of available resources, it will endeavour to provide a selection of the most suitable options.

One of the key underlying concepts is one not often considered within security due to its rare necessity. As the focus has shifted towards biometric environments, it has only been a matter of time until user well-being will become less of an abstract concept and more a practical consideration. This is a natural progression as when a sample has been verified as living, by liveness detection, then the next stage is to make sure that the sample provided is done by the user's own free will. As with many coercion detection concepts, this area has not been considered in any detail, due to the emphasis on liveness detection as a more important security sub-system. Therefore, this area is very new as there are only very minor references in certain works such as [17] [86] [87]. These discuss very briefly the concept of coercion detection as an area of importance and [17] identifies some very basic techniques of coercion detection, alongside some very general headings in which to work within, as identified below.

Consequentially whilst current research has only identified some general techniques there potentially many more that can be used to detect coercion ranging from the simple, e.g. using an external monitor 'panic key', to the complex "affects detection and analysis"which can cover a variety of chemical, biological, and behavioural standards. As with all factors related

to biometrics there are many considerations that can improve or detract from the success of a specific technique such as accuracy, multi-modal fusion and technological heterogeneity.

Therefore the next section will cover what characteristics coercion detection has and how to develop a number of techniques for coercion detection.

3.5.2 Coercion Detection Characteristics

Whilst there are a number of ways to gather coercion samples, mostly consisting of physiological characteristics associated with a variety of emotions, there are techniques that can be developed that rely on the emotions themselves.

The main characteristics to use for coercion detection is fear, but other similar emotions such as stress and disgust also have places in these techniques. Considering fear as a primary example, as other emotions can be measured in much the same manner, there are a variety of different techniques in which to gather the data, most of which are triggered by a variety of biological signals within the body. These can be gathered at a very high level as [179] postulates saying that when using functional magnetic resonance imaging healthy humans, required a negative connectivity with the cortical and sub-cortical pathways towards the amygdala (set of neurons within the temporal lobe) therefore potentially enabling fear to be detected. However, whilst it may be possible to detect fear in this manner, the technique required would prove prohibitive within most installations, except in extreme circumstances, due to the cost, implementation difficulties and acceptance of the technique.

Alternative techniques normally draw upon stress as an indicator of fear. However, this has a host of potential problems. To begin with stress is widely considered as the most ubiquitous health concerns within the modern workplace environment therefore bring to question its suitability for coercion detection [24]. This leads to questions such as how can "normal"stress be identified from stress caused from coercion. This emotion based evidence is commonly known as affect, or affective feedback and has been studied for many years, primarily within the HCI and machine learning areas as identified by [146]. Recognizing affective feedback is important for intelligent human-computer interaction and [123] indicates that techniques for machine learning to detect emotions from the human users need to develop a more predictive and usable system. Therefore, whilst this technique may well have valid applicability within coercion detection, it must be considered with regards to where the installation is happening, as the location and environment will denote the influencing factors. Consideration must be given to the installation location and therefore may encounter a huge variation in levels of stress, and therefore a potential problem for fear detection. For example, generally speaking, a home environment would be less stressful than a work environment,

therefore a technique may differ in relevance, different workplaces may have different levels of both standard and peak levels of stress [66].

These factors underline some of the basic concepts within coercion and the current thinking is to split technique by their individual requirements, in much the same way liveness detection characteristics are considered. These categories are involuntary, voluntary and environmental. Therefore it is these areas that will be considered first.

Involuntary approaches

Involuntary approaches to coercion detection include a huge range of factors that cover techniques that the user has no control over and are direct responses to stimuli. Coercion detection requires a repeatable emotional output for example, fear. This output can then be correlated and an authentication can occur accordingly depending on the relevant level of the technique. One universal problem with these techniques is the range of potential factors that can cause the physiological effect to occur, this noise data is of huge importance, and adequate noise cancellation techniques must be developed to produce accurate techniques. One such technique would be to incorporate an autonomic factor within a system, such as context awareness that would allow for a dynamic target data range to be set depending on environmental factors. Without techniques such as this, noise can have a huge effect, for example, if an elevated blood pressure which is used as an indicator of fear due to the fight or flight reflex which causes an increase in blood pressure as identified in the seminal work by [15] and iterated upon for the subsequent Cannon Bard Theory [132]. Then how can this be differentiated from simple blood pressure elevation caused by medical noise or physical excursion. Therefore, when developing with these techniques and consideration the onus must be put on technique and its capabilities to avoid any potential problems within this theory.

Physiological signals, whilst potentially very effective, are gathered from different areas and differ somewhat in style, for example, some are based around noise deviation, whilst others are medically based. Traditional noise deviation causes security samples to be refused as the sample has been changed beyond suitable levels. Examples include changes in voice pitch, often an issue when dealing with samples that are gathered within areas that have large amount of excess noise such as background sound, environmental effects etc. These noise deviation techniques can be very applicable, however, they suffer from a high degree of intra-user variance [17].

Whilst noise and medical deviations are often very similar, there are some differences, and whilst each can affect the other the separation provides a more robust description of the factors involved. Medical data techniques are reliant on medical fluctuations to denote

coercion such as an increase in heart rate, perspiration etc. most of which are often also used to indicate relevant emotions such as fear, happiness etc. The medical based indices have the potential to be very transient, depending solely on the comparative nature of the users, being susceptible to illness, as some may change the range of acceptance for coercion detection beyond normal levels.

Whilst it is often considered that medical based data is purely physiological constraints, [158] show that there is a specific link between stressful experiences and psychological factors. Therefore a selection of involuntary speech pattern can be considered, how speech changes when under stress, coercion etc. However, this can be difficult to incorporate due to the sheer amount of noise variance from all factors of biometric security, and if voice techniques were implemented then they would have to have a robust degree of fusion.

Voluntary Approach

Whilst most coercion techniques revolve around physiological data, in some form there are others that can be developed that may be as relevant and applicable depending on the scenario. These techniques are voluntary approaches because they rely on the user to provide data of some kind, and whilst it can be argued that medical techniques are also voluntary in nature, the distinction is that there is additional conscious thought to provide a sample. This is normally done with non-medical data, for example, speaking a password or selecting a pattern.

There are a variety of forms this could take. For example the user may carry a key which may be used as a 'panic alarm'. A second approach would be to utilise a selection of pass-codes/keys that would denote coercion. Although to be thoroughly effective the technique would have to be completely integrated within the full biometric system, otherwise the use of the traditional key system would be seen as a simple and easier approach to security. Without the thorough integration it would also become obvious to the attacker that some preventative measure is being taken, as the inclusion of a key in the final stage would become suspicious for a nefarious user, and therefore reduce the effectiveness of the technique.

Environmental Approaches

Whilst the other techniques focus on the user as a data provide, environmental techniques instead utilises the environment the user is within and the appropriate responses therein. This technique can be problematic to implement as it does not focus on the different variants of the user, instead it correlates data regarding the user as a focus point. For example the use of cameras and proximity maps can depict if there are people close together. Whilst on its

own, this is both irrelevant, and can easily be explained away with a handful of reasons when combined with appropriate security protocols the development can become more impressive. For example, all users must make sure they are on their own when authenticating into a system, no more than one user with a biometric scanner at any one time etc. Therefore, if a proximity sensor is able to detect multiple people closer together then it can indicate an attack of some kind.

Obviously there is a host of potential problems with this form of approach, least of all the ease of misunderstanding. Using the above example, if two users were carrying a heavy parcel, they would be identified as too close and therefore the system would respond, in this case erroneously. There are also other noise based concerns, such as how other data sets impact the system, can humans be identified specifically or will other species cause a false rejection, will other heat signatures set of proximity response and if so how can this be dealt with. These false reject samples would create a generally untenable system and it is for this reason that most environmental techniques have been dismissed in this research.

3.5.3 Coercion Characteristics

Whilst the discussion and development of actual techniques is very important, there is a plethora of other concerns such as the identification of the underlying theories relevant to the development and implementation of coercion techniques. The intent is to adapt the same factors considered for liveness detection so that it is applicable to coercion. This section will focus on the identification of important coercion concepts, and not the development of individual techniques, which will be covered in a later chapter.

Performance

The performance of a computing system is often the most important factor and it can be one of the hardest to identify due to the potential range of factors that can contribute to it. Within coercion detection performance is no less important and denotes the success rate of the detection process, along with the comparative ease the process was conducted in.

Due to the potential variety of coercion detection the performance of the individual methods can differ dramatically depending on the type and effect of said technique. For example a tangible 'panic keys' can be have a high performance measurement as they are easy to use, cheap to manufacture and already quite ubiquitous in society but can be easily circumvented. Alternatively, techniques such as affect testing are much harder areas to identify and require a more thorough technology integration to provide efficient installations. This is primarily due to the lack of specific research that focuses on using these techniques as

coercion detection concepts instead of just affect/emotion detection. Therefore the reliance on novel technique development such as techniques based on FACS [54]. Most of this research, into stress and lying detection, has been conducted in other areas such as [158]'s work based around poker, and whilst the focus is different to coercion detection, the premise postulated therein can still be applied due to the underlying non-specificity. For example, this particular work identifies that detecting stress and lying only achieved an 82% and 71% success rate, which would provide a very poor degree of security, however this figure would have to be taken into account considering the age of the research, the subject area it is being considered in and the specific techniques of data collection as other research has had marked superior results such as the 90% success rate within [146]'s work. Another flaw, for direct security use is, that this technique utilised long periods of testing, where testers would have data collected about them for an extended period (approximately 15 minutes) which would not be applicable within coercion detection, as like biometric security and liveness detection the speed in which the sample is dealt with is of utmost importance and a direct efficiency factor.

Whilst these techniques could be utilised in a variety of situations the exact implementation would be colossally ineffective due to the time taken to gather samples. Therefore any techniques would have to be suitably adapted to the needs of coercion detection and the specific environment being designed for. The main point of these distinctions is to show that the effectiveness of a technique is not limited just to the current lack of techniques, instead it is only limited by an overall lack of research and understanding of the scenarios it is being utilised in.

Heterogeneity

Heterogeneity, whilst often focusing on devices, is a factor that can have long reaching implications throughout the development of system. Within coercion detection the heterogeneity of techniques can also be considered as the specificity level. This details the ease in which the techniques can be incorporated across multiple scenarios, techniques and devices. Therefore, it must be identified, as thoroughly as possible, what techniques are most relevant and which are best suited to individual implementation.

As coercion detection is an extension of biometric security the comparative link between security, liveness and coercion must also be considered and this tri-modal will identify what is most relevant within the situation. A lack of heterogeneity here will cause numerous problems to occur such as lack of efficiency. These factors lead to the following assumption, that the most effective techniques will be those that can be used throughout the different

levels of biometric security. Therefore, creating a thoroughly integrated multi-modal, and multi-security system.

General Fusion

When dealing with fusion the main focus is to combine different techniques, from whatever stage of the security process they are at, in the most efficient and effective way possible. This is normally done within each individual section, for example a multi-modal biometric system would contain iris and facial recognition to improve the overall degree of security, within liveness detection blood pressure and skin conductivity tests may be undertaken. Whilst this is the normal route, there is also the additional concept of multi-layer fusion.

Coercion detection has a number of fusion considerations and is perfectly viable within a biometric system. Coercion fusion, like other system, would have to consider how they integrate together, as well as how multi-layer integration could occur for example with liveness detection. An example of this would be Figure 3.6. Whilst it is possible to have three completely separate layers, the thoroughly integrated system could provide a greater degree of security, and protection against security threats. There are factors to consider such as threat vectors, liveness detection styles, etc. as failing to do so would produce inadequate fusion and potentially a reduction in security potential.

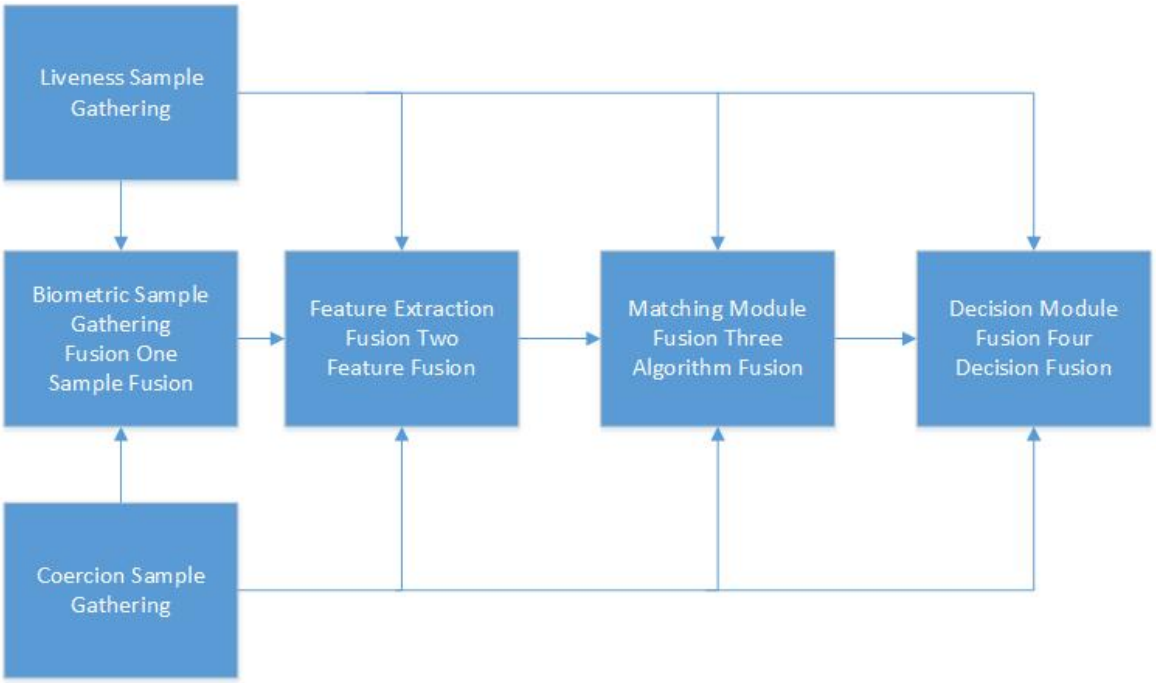


Fig. 3.6 Potential fusion areas

The incorporation of coercion fusion into multi-modal systems would be a bigger challenge as there would be an additional level of fusion to consider as demonstrated by Figure 3.6. As well as the initial fusion between coercion techniques, layered fusion would have to be considered within multi-modal systems. Fusion that would contend with security, liveness, and coercion methods would have to consider the relevant techniques in more detail, as whilst one technique may be acceptable for a security and liveness environment, the incorporation of coercion may turn the technique into an ineffective style. For example [70] identifies that when utilising multi-modal vocal and emotive systems the effectiveness drops with emotive speech, therefore showing that fusion does not necessarily mean automatic improvement. Alongside this would be specific extra problems, as the emotive aspects would be utilised as an integral part of coercion detection due to the emotional undertones of the subject, therefore degrading the technique even further. This exemplifies the necessity to make sure that the techniques used are the most applicable, and whilst this is very difficult to do currently due to the lack of techniques and the poor categorisation system,

Cultural implication

Unlike many areas of computer science, where the effect of culture within the subject is minimal, coercion detection has some definite cultural implications. These occur due to the innate individuality of coercion detecting techniques, and it is not just cultural but individual implications that can dramatically effect the way coercion detecting techniques work. For example, certain cultures may find specific techniques distasteful and therefore there would be an unwillingness to accept their integration into systems [123].

Medical Implications

Medical data comprises the main bulk of biometric detection therefore demanding constant reinforcement of validity, permanence etc. These factors can be dramatically affected by medical intra-variance due to the ever-changing illness or medication effects. These effects can dramatically change a sample, to such a degree that they would classify as being coerced even when they are not, completely due to the medical noise factors. For example a user that is being coerced may have an elevated blood pressure due to the physiological, and psychological stimuli identified in the fight and flight reflex postulated by [15] and discussed in countless work such as [58]. Alternatively the user could have jogged to work similarly elevating their blood pressure. Whilst this concept is overall simplified the premise holds true throughout the technique. The premise of this concept is that a comparatively simple external influence, which is not the expected response, provides the data that would be associated

with coercion, consequentially bypassing the system security. This highlights the potential consideration as there are more complex intra-variations, deviations that may affect the sample. Following on there are a host of medical conditions that can change a user's blood pressure dramatically, therefore providing false readings to be gathered, alongside these are a variety of samples altering considerations, all of which may become a security risk. Some of these techniques are very ubiquitous such as stress related blood pressure increase [58] to the rarer, but more serious condition, such as strokes [99].

Therefore, it is imperative that the different techniques are assessed on the relevant susceptibility to medical factors, a factor that becomes even more important when considering the fusion characteristics, as the combinations of similar techniques for both liveness and coercion, whilst potentially efficient, would provide a large target for nefarious user's spoof attacks based on medical data. Additionally, if the same characteristic for liveness and coercion detection are being used as a standard, then they are even more so susceptible to medical noise that affects the singular data type, for example blood pressure deviations.

Whilst the focus for coercion detection is security, and the medical intra variations are an obstacle to overcome, there are other factors that can be integrated. As with liveness detection and security, there is the potential to gather the medical data when sample collection occurs, as the sample is in truth a piece of medical data that can be used for medical analysis and a host of other factors. This data can potentially be used for a number of other technologies, and whilst the ethical and legal implications are many and varied, the focus within this section is in the collection. The intent is to minimise the impact the legal and ethical consideration have, instead this work will focus only on the technological issues, leaving the ethical etc. to another work.

When samples are being collected, the user is providing medical information, often very specific medical data such as blood pressure, salinity levels, ECG and EEG factors (obviously depending on hardware capabilities). This leads to a variety of potential options, for example, if the data, gathered from the biometric system is stored, then a time based map could be generated for the user's personal records, or the inclusion of health professionals, depending on user choice. This data could then be used to provide real-time health monitoring, as well as historical health monitoring. Alongside this, incorporation into a smart system would enable user profiling techniques to be deployed. These context aware environments would be able to adapt to the user's needs and profiles utilising an SOA environment. Whilst this represents a pervasive environment, focusing on user services, there are other techniques that could be included. For example the inclusion of these medical profiles would allow the creation of dynamic security systems, that would evolve with the user and as the users' health changes the relevant acceptance metrics could change alongside. This would also enable

recurring deviation to be identified and adapted to, for example, if a user engages in vigorous exercise at a specific time each day/week etc. and normally this exercise interferes within the authentication, then the system would have the capabilities to identify these factors and accordingly adapt to facilitate them. Whilst this is a generalisation it has a lot of potential, however there are considerations that would prevent it from working optimally. As this factor would have to occur with a degree of regularity the process would be the same as normal authentication, and therefore would provide no different security threat level, unless it can be shown that higher data ranges are harder to authenticate [123].

User Acceptance

A traditional problem with new technology, and especially biometric systems is the user acceptance of the technology [82] [81]. Whilst users can sometimes be reluctant to adopt new technology, for a variety of reasons, including age, background, opinion etc. [56]. As [114] identifies biometric samples deriving from data that is often very personal to a user, data that is rarely called upon for any other reason, except for a medical situation, or unwanted situations, immediately providing cause for potential consternation. For example the taking of fingerprints may well have criminal connotations for users due to the technology's ubiquity within law enforcement scenarios. Liveness samples may include blood pressure monitors which obviously have a number of definitive health connotations, including the user's unwillingness to accept medical information, to embarrassment etc. The same features can also have the same problems for coercion detection. Therefore, when implementations occur, sufficient consideration must be given to the acceptance of the technology, because if a user does not wish to use the device/technique, then it is worthless.

This problem is highlighted when considering two factors, the first identifies that the techniques are medical in nature, and many people are reluctant to discuss/provide medical data, especially if the data is not being used for medical purposes. This reluctance would cause a host of issues for most biometric, liveness and coercion techniques as the primary form of sample would be more difficult to identify and utilise. Secondly biometric systems have had a very bad press within the media and there are numerous occasions where biometric systems have been spoofed due to the loss of or theft of a sample, both within real environments, and a host of popular cultural formats. Whilst these opinions are often erroneous, due to the elaboration of the media, the poor public opinion has sustained and therefore further alienates users.

This problem is one of the more difficult to address as it deals with changing the opinions of users, something that traditionally is very difficult. To achieve this there are many different techniques including a thorough education process, identifying that biometric environments

are as safe, and not as portrayed within popular media alongside a progressive and systematic inclusion of the technology within popular mediums. For example the inclusion of biometric security options within smart devices will allow the users to become familiar with the technology.

Noise

Many of the factors that are relevant for the noise within biometric and liveness detection will also be relevant for coercion. However, there are some subtle differences as coercion is not authenticating or verifying a user. Simply identifying if the user is being coerced or not. Therefore, whilst the effect of noise is important, it will not have the same degree of impact as it does within the initial biometric security processes. Whilst it may not have the same degree of importance as biometric security it is still needed so that the exact effect noise has on coercion detection is understood. The traditional understanding is that noise provides extra data that detracts from the overall sample, therefore making it harder to authenticate.

As most coercion detection techniques identified within Table 3.5 are based around medical data and it has been shown that medical noise can exist from a variety of sources such as exertion, medical condition etc. The effect of this will be of utmost importance to research. This noise will make the sample deviate somewhat from the expected range, therefore potentially preventing coercion acceptance. One theory could include the integration of robust coercion detection algorithms that attempt to remove any noise that occurs, this is something to consider for future research. The identification of what form the noise data takes is something that must be considered. Is the noise medical in nature as discussed earlier, if so what are the potential proofs against it and how can it be dealt with? Are the factors, environmental, changing the sample enough to impact the authentication process, if so how susceptible is the technique to these factors etc.? These factors must be taken into account when constructing coercion detection techniques as without these factors being addressed the more chance that a security threat will occur and if this threat can exist then the overall effect of the system will be reduced accordingly. Therefore the next stage in this development will be to identify suitable coercion detection techniques, whilst there are some basic techniques discussed in the limited research currently conducted in this area, the emphasis will be placed on the novel techniques that have been developed for this work. These techniques are based around current ideas which have relevance to coercion detection, however, it is rare that they are identified as coercion techniques, instead they are developed from ideas and research that functions within the area, which can be adapted to coercion detection, with careful understanding.

Technique	Voluntary Involuntary Environmental	Core concepts	Medical Noise	Tech Difficulty	Other Noise
Voice pitch variation [158] [17] [51]	Involuntary Voluntary	Humans are often poor at detecting affect (emotion) due to culture, experience etc. Can be either involuntary and voluntary	High	Low	High
Skin conductivity peaks [158]	Involuntary	Increased conductivity	Medium	Medium	High
Heart rate variability (HRV) [17] [158] [72]	Involuntary	Look at fear notes for actual data about the HRV.Overall poor due to the time to take data (iCalm = 15mins)	High	Medium	Low
Language differentiation [17]	Involuntary Voluntary	Clue based around language pattern, slang, dramatic change in style of language etc.[51][70]	Low	Low	Medium
Verbal keys [17]	Voluntary	Dialogue features, word usages, passcodes etc. [123] Dependent on memory of phrases, codes etc.	Low	Low	Low

Tangible Keys [17]	Voluntary	Conscious key solutions, wearable computing, internet of things etc. [122]. Pervasive inclusion, how can pervasive work be linked into everything. Use of panic buttons is well established, and if incorporated into mobile devices, or a "emergency pattern" for pattern keys etc.	Low	Medium – potentially high	Low
Intended false authentication technique [17]	Voluntary	Providing deliberate incorrect finger for recognition, knowing that it will be caught. Iris/retina/facial scanning allowing user to look at "fear zone" if user looks there for specific amount of time it counts as coercion	Low	Low-Medium	High
Facial Movement[17]	Involuntary	How robust must facial movements be, look at facial action coding systems by. FAC system is important by Ekman and Friesen, good for higher emotion, but others say Bayesian networks are more robust technique	Low-Medium	Unknown	Medium
Body Posture measurement [17]	Involuntary	Body posture measurement – detecting from posture – how viable, when initially authenticating may not be that valid, within a continuous authentication environment[70].	Medium	Medium	Medium

Emotion detection [17]	Emotion	Multi-area – detect number of emotions fear, stress etc. Subject-elicited versus event-elicited: Does subject purposefully elicit emotion or is it elicited by a stimulus or situation outside the subject's efforts?Lab setting versus real-world: Is subject in a lab or in a special room that is not their usual environment?Expression versus feeling: Is the emphasis on external expression or on internal feeling?Open-recording versus hidden-recording: Does subject know that anything is being recorded?Emotion-purpose versus other-purpose: Does subject know that the experiment is about emotion?	High	High	Low
Tear Detection [17]	Involuntary	Crying detection, Is it possible to detect, can it be watched for.	Unknown	Unknown	Unknown
Fidgeting and Micro Movements [17]	Involuntary	Fidgeting, and micro movements would cause general authentication to be problematic, this could be an automatic detector, again problem with resources allocation. Most of the best classifiers are still only 70% accurate. [70]	Medium	Medium	Low

Sensor Based solutions [17]	Environmental	Video cameras that can detect the proximity of users to each other, potential for coercion detection. Can us proximity, heat signature etc.	Low	Low-Medium	High

Table 3.5 Coercion Detection Techniques

3.6 Biometric Conclusion

The plethora of biometric information gathered and disseminate has helped identify a host of biometric considerations that are of the utmost importance such as the method of categorising liveness and coercion techniques along with some of innate problems associated within liveness and coercion detection. The main area of consternation is the lack of uniform techniques to compare liveness detection techniques.In addition to traditional authentication techniques most biometrics can provide additional data that could potentially be used for other system features such as: medical information providing real-time health monitoring, identifying suitable biometric techniques for users that may have specific needs that render other techniques invalid, etc. This highlights the potential advantages this dynamic biometric environment would provide a security installation alongside other areas such as smart homes/ihome/pervasive, which can be used for a medical purpose in the care and monitoring of certain illnesses and disabilities.

Currently, there are many researchers that use their own metrics to measure similar factors such as FAR and FRR which makes it very difficult to compare accurately liveness techniques. This lack of comparison makes developing fusion based systems problematic along with identifying suitable areas for research also indicating how categorising these research areas is key to improve the understanding of the area. By correctly categorising different liveness techniques, informed decisions could be made to highlight what areas are most useful to research, as well as minimising a number of superfluous data collections metrics and measurements within the overall research area.

Currently coercion detection has been identified as a potential area to consider how there has been limited research conducted into this area. This has highlighted some basic ideas such as the three types of coercion technique (voluntary, involuntary and environmental) and the concept of using psychological factors as methods of detection. It does not identify how this will occur and what factors will effect the development and implementation of coercion techniques. This is an area that needs to be researched, if coercion detection will become an important factor in biometric security.

Chapter 4

Taxonomy Development and Application

> They say a little knowledge is a dangerous thing, but it's not one half so bad as a lot of ignorance
>
> Pratchett, 2004

One thing that has become apparent is the necessity to be able to classify different liveness techniques as there are problems occurring due to a lack of standardisation within the research. This can be solved by creating a taxonomy that allows data to be standardised and more easily compared. Without this technique it will become increasingly hard to develop liveness fusion techniques and research different liveness areas. Therefore a method to discern the most suitable device or technique is of paramount importance to better understand the area. This would also provide a method of technique or device comparison so that a developer/researcher would be able to choose the most appropriate for their needs. This is especially important within biometric security as biometric techniques have a huge variety of characteristics that can differ dramatically depending on the techniques, device and installation, highlighting the importance of cohesive categorisation. Therefore the development of the taxonomy, intends to classify all of these techniques in a manner that can be used by researchers, developers or even autonomous systems to choose the most suitable device and techniques for the scenario. Whilst the initial classifiers used were concerned only with biometric devices, it became clear that liveness detection was as necessary to classify which subsequently led to coercion detection and its need for classification.

The first question to consider is why use a taxonomy? The main advantages of a taxonomy is that it allows the viewer to see the data when compared to similar data sets within the area, Instead of only seeing one particular data range. It is also a systematic method of data gathering and classification. Whilst classification is normally done to group pieces of data together, to allow a more easily followed analysis, the taxonomy will also try to categorise

these groups, name them enabling a critical review of their suitability. Whilst in the future it is possible that this taxonomy might develop enough that it becomes an ontology, which would become much more in-depth, and would develop a much more rigid relationship identification between different areas, currently due to the comparatively small quantity of data, and the specialist area it fits the taxonomy will be used.

To this end the taxonomy must allow both techniques to be identified, compared, and analysed whilst also providing the framework to allow for future innovations to be included, this scalability is needed as the biometric security area is rapidly changing and any taxonomy that is static would not be usable.Therefore this chapter highlights this taxonomy creation alongside the appropriate steps therein, the developing of liveness categorisation alongside the development of the coercion detection standard.

4.1 Liveness Detection Development

One of the most reliable techniques to identify suitable devices,solutions to problems, etc., is to categorise the available techniques and choose the most relevant for a scenario. Within security this process is even more important as the suitability of a security technique can mean a successful defended system instead of a breached system. As biometric security techniques contain a huge variety of characteristics, which can differ dramatically depending on the technique or device, the importance of choosing the correct security features is vital to the health of aa secure system. These characteristics, as discussed in chapter 3 cover the main features of a biometric device and must be considered when developing a multi-modal biometric systems. This is because the use of multiple devices leads to biometric fusion which must take into account the different characteristics within each device and failure to do so will create a system with inadequate fusion capabilities which will lead to an inferior security coverage. This is demonstrated by [87] who compared PAD (presentation attack detection) techniques with comparison subsystems (security matching), and found that this fusion technique did not improve the accuracy and efficiency, which is the main focus of fusion systems. In fact the FRR increased from 0.58% to 3.55% therefore showing more false reject rates within the system therefore reducing the overall efficiency. This incorrect choice of techniques would be preventably by using the following taxonomy therefore allowing a more suitable combination of techniques to be identified and implemented.

Table 4.1 identifies the current classifiers of biometric devices, this basic taxonomy shows the six classifiers which highlight the strengths and weakness of biometric devices. Whilst this technique provides some valid information it uses an ordinal measuring technique [90] which does not provide all the data that is required such as difference between 'High' and

Biometric	Universality	Uniqueness	Permanence	Collectability	Performance	Acceptability	Circumvention
Face	H	L	M	H	L	H	L
Fingerprint	M	M	H	M	H	M	H
Hand Geometry	M	H	M	H	M	M	M
Keystroke Dynamics	L	M	L	M	L	M	M
Hand Vein	M	H	M	M	M	M	H
Iris	H	H	H	M	H	L	H
Retina	h	L	M	L	H	L	H
Signature	l	L	L	H	L	H	L
Voice	M	H	L	M	L	H	L

Table 4.1 Initial Biometric Classifications [82] [7]

'Medium'. It is for this reason that the proposed taxonomy will use a interval measuring system which is discussed further in section 4.3.

These characteristics are important when deciding on biometric devices as they provide interested parties the most relevant features and techniques to be chosen, therefore improving the potential suitability. For example if a system demands a biometric with a high degree of universality and collectability then facial recognition will be a good choice due to its high universality characteristic biometric. However it is still very susceptible to noise data such as light, facial occlusion etc. Even when considering high universality devices there are always problems to face such as people who suffer from Treacher Collins Syndrome, a genetic disorder that effects the development of bones and tissues in the face [77]. This would prevent the use of facial devices and could impair other facial based techniques such as iris/retina scans.

Whilst these characteristics are commonly used within biometric devices: liveness detection has no such categorisation, in fact there is very little linkage between different forms of liveness detection mainly due to the variety of techniques available to different devices. Therefore to improve the understanding and access to liveness techniques this taxonomy will be proposed that will improve access and fusion understanding of researchers and developers.

Taxonomies are widespread within computing, however there are underused within liveness detection as identified within [107]. This details the different levels of biometric devices and sub-genres of biometric styles detailed within [184] which emphasis behavioural biometric categorisation. This highlights that there is no current general taxonomy, due to the comparatively newness of the area, a taxonomy would provide a variety of benefits such as:

1. Liveness detection techniques are linked intrinsically to their biometric parents therefore should the categorisation process follow the classifiers used by biometric devices; should a completely new set of categories be developed; should a taxonomy developed that uses the most appropriate features of the biometric devices.

2. Biometric devices utilise the sample data only, therefore how should the liveness data be gathered? There are three methods that could be used:

   Intrinsic - Automatically taken as part of the initial sample

   Innate - Can be gathered with the initial sample but not automatically taken

   Voluntary - User provides additional sample

3. How does biometric fusion affect this characteristic is liveness fusion affected by the same factors as biometric fusion .

Therefore the following section will develop and analysis the suitable classifiers for the taxonomy, highlighting their suitability for liveness detection and how they will contribute to the overall algorithm. This will then be followed by the overall taxonomy construction and evaluation.

4.2 Liveness Detection Categorisation

Biometric devices currently use a basic taxonomy that considers six different factors and whilst this works it is not optimal. There are two main problems when categorising things.

1. What will the categorisation actually provide, will it benefit the area or will it merely change the 'aesthetic' of the data organisation.

   To address this the characteristics have been analysed to demonstrate the potential impact on liveness techniques. These characteristics will provide a broad range of information to a developer or researcher, therefore enabling them to choose the most suitable biometric liveness techniques for their system. This point is moot if Uni-biometrics are being used, however in this taxonomy uni-biometric systems will not be considered as [22] and [82] identified that uni-biometrics are very rarely used due to their inherent weakness, and therefore this categorisation process will focus solely on multi-modal biometrics.

2. The second problem is that there is a lack of liveness detection standardisation as most research has been focused on specific techniques such as [80]'s identification of

fingerprint techniques, so little consideration has been given to liveness detection in multi-modal biometric systems [121]. Therefore new considerations have appeared when choosing biometric fusion devices. The developer must take into account how liveness detection will affect the multi-modal system [114]. Therefore the taxonomy will provide a comparative ranking system that will allow researchers and developers to better choose the most suitable techniques for their systems.

This taxonomy will be developed following a top down technique which will identify the important factors within the liveness detection endeavouring to keep the overall taxonomy broad and shallow therefore keeping it as simple and elegant as possible to prevent unneeded complexity [34] [124]. This follows the inductive methodology being used (Grounded Theory) as it uses observations on the current available data allowing to help mould the taxonomy constitute parts. Therefore whilst the development of a completely new taxonomy is a possibility it is not the most efficient technique as there are a number of taxonomic elements already available within the biometric device taxonomy identified by[81]. Whilst these characteristics can help develop a base line there are specific problems that need to be covered allowing the integration of liveness and coercion detection, which it currently cannot do. This will be done by utilising the current characteristics as well as endeavouring to maintain their robust and applicable nature, whilst also providing flexibility. It is for this reason the top down approach is being used [124].

There are a number of characteristics within liveness detection that have been inherited from biometric devices and whilst some features are relevant others are less so. Some of the current biometric characteristics identified may well work with some minor re-adjustments however others are not quite as pertinent to liveness nor coercion detection which is discussed later in this chapter. Therefore this is the route that has been chosen because there is a rich research resources available regarding the taxonomies of biometric techniques and devices, such as within the works of [147] [155] [148].

Table 4.2 identifies a cross section of liveness detection categorisation requirements deemed necessary by a host of respected researchers within the field [166] [121] [133] [81]. Whilst these sources identify a number of potential categories for liveness detection, they do not focus on this area instead identifying the need for the requirements whilst encountering security threats relevant to the biometric device. Therefore there are numerous areas of crossover between the different researchers, there is very little collective comparison, subsequently this is what was identified as the first stage, the data assimilation and analysis. Table 4.2 represents this raw category identification, especially as they are deemed relevant, because many of which have roots within the original categorisation of biometric devices as identified and developed by [100] [114].

Characteristics	Toth [ ]	Schuckers [ ]	Tan and Schuckers	Pan et al [ ]	Reddy Et al [ ]
Ease of Capture	Yes				Yes
Available within sample (Intrinsic etc.)	Yes		Yes		
Available within techniques (pulse in iris scan etc.)	Yes		Yes		
Additional Hardware/Software needed	Yes		Yes	Yes*2	Yes
Universality (no pulse in iris scan etc.)	Yes				
Biometric Heterogeneity (how easily can it be used across biometric techniques)	Yes		Yes		
Independent-ability (How much additional input does the user need to include i.e. challenge response etc.	Yes		Yes		
Accuracy of the liveness technique (is there a liveness variance of the FRR/FAR etc.)	Yes		Yes		
Permanence (How mutable is the sample – issues surrounding disability/illness and biometrics.)	Yes				Yes
Combinations (How easy is it to combine the same technique across multiple biometric techniques e.g. fingerprint + hand print + vein scan etc.)	Yes				
Specialism (Is it only suitable for one technique, if so is it more precise etc.)	Yes		Yes	Yes	
Ease of capture	Yes				

Permanence (Does the liveness sample differ with age/illness/disability etc.)	Yes				Yes
Ease of implementation and use (Will it cause users any adverse effects	Yes		Yes	Yes	
Does the gathering technique effect the collectability of the technique	Yes		Yes		
Ease of Collectability	Yes	Yes			
Cost	Yes			Yes	Yes
Robustness to attack (Saline solutions, algorithmic integrity etc.)		Yes	Yes	Yes	Yes
Transparency (User does not notice the system, if having to provide challenge response etc. then it is not transparent)		Yes	Yes	Yes	
Fusion capabilities/flexibility			Yes	Yes	
Technology flexibility (e.g. available SDK)			Yes	Yes	
Flexibility (Technical using different algorithms etc. for different features)			Yes		Yes
Noise Effectiveness (Is it susceptible to noise, if so does it have any counter measures)				Yes	
Ease of spoof data collection					Yes
Can the data be used for other purposes (medical smart homes etc.)					Yes

Universality		Yes		
Uniqueness	Yes		Yes	
Permanence		Yes		
Collectability				Yes
Performance		Yes		
Acceptability	Yes			
Circumvention			Yes	

Table 4.2 Liveness Detection Characteristics

4.3 Liveness Detection Analysis

The current biometric taxonomy consists of seven categories covering the salient strengths and weaknesses of biometric devices, this process is highly useful to developers and researches when device inclusion is being considered. However, whilst this technique has been used for a number of years, some problems have become apparent. The main problem is that the technique used to measure biometric device characteristics is based on an ordinal system [90]. This only contains the factors: High, Medium and Low, and whilst this system has been acceptable, primarily due to the limited base factors being measured, it is not an effective measuring style. This is because ordinal systems are unable to differentiate between the different levels, therefore whilst it is possible to identify that 'High' is a greater value than 'Medium', it is not possible to identify in what way or by how much. This can be a major problem as systems depend on very tight efficiency and low margins of errors to function and this ambiguous metric can cause a host of problems that would detract from the overall efficiency. This is highlighted within the work conducted by [22] utilising iris and fingerprint fusion systems. The evidence showed that fusion only improved when the impact of both techniques was changed dramatically, and secondly that multi-modal environments were more robust than uni-modal environment. These factors show the limitation of the ordinal system as these result are well known within the research community as highlighted by [114] [25][40] [139].This research would have benefited immensely if the taxonomy had been developed using an interval measuring system, which highlights the statistical difference between different levels. This would have allowed researchers to choose similar techniques due to their interval rating instead of having to rely on the original ordinal system. This would therefore provide a more thorough understanding regarding the fusion requirements of different techniques, something the current ordinal system is poor at providing.

The method that this research takes is to change the measurement type into something more usable for the users [167]. Currently there is no differentiation between levels subsequently any correlation between the levels is impossible to postulate, for example is 'High' twice as good as 'Medium' or is it three times as good?. Whilst this technique of data description is excellent in its place, such as when it is not possible to know the relevant statistical difference between one variable and the next, within the biometric characterisation it is possible to identify the difference between measurements such as the difference between device's FAR/FRR/FTE and other mathematical data such as the median deviation of impairments that can effect a specific device. For example according to [29] there are about 1.7 million amputees living in the U.S.A, 65,000 of which included one or more fingers. When considering this in line with the population of the U.S.A in 2013, 316,148,990, this means that when measured to two decimal points 0.02% of the population would be unable to use a

fingerprint scanner, and whilst this is a generalisation it illustrates the problem with an ordinal measuring system. When also considering iris recognition problems which include factors such as Anophthalmia and Microphthalmia [73], which affects one in every ten thousand live births, which would equate to 0.01% of the population being affected, the indication is that iris would have a having a higher degree of universality. This allows an interval value to be developed as it can be said that, in this example, iris recognition is twice as universal as fingerprint techniques. This interval system is being developed to allow more precise data to be developed that will allow developers and researchers better understanding of the factors within their own work.

To preserve the highest degree of familiarity, and to avoid unnecessary confusion following the broad and shallow technique, the basis for this categorisation process begins with an analysis of the traditional biometric security categories and their relevance to liveness and coercion. One issue to contend with: some of the inherent problems within the device taxonomy will become similarly problematic for liveness and coercion and whilst this is important, the emphasis must be placed on the evaluation and development of the adapted taxonomy and will not dwell excessively on the problems within biometric device taxonomy. It must be understood, that whilst the technique needs to be as efficient as possible, removing unneeded factors, if the inclusion of additional factors is deemed pertinent then it must be undertaken.

One factor must always be considered: is feasible to keep the same terminology as the similarity of terms may cause confusion within security systems. This could cause confusion for system designers or researchers however it would allow a greater degree of fusion and heterogeneity. Alternatively the categorisation process would have to be made more dissimilar to avoid confusion but it would potentially increase the complexity of the taxonomy. Whilst this is a valid consideration, this research will focus on making the categorisation process for liveness and coercion detection as similar as possible promoting a heterogeneous environment that will improve the potential fusion between devices and techniques.

The initial stage of development was to remove redundant or repeated characteristics within the sample data, whilst this was not an extensive process it did produce a richer data set. The following stage was to combine the remaining characteristics therefore limiting repetition and redundancy. As has already been discussed, the starting place for the taxonomy development is within the biometric characteristics.

4.3.1 Universality

Universality is as important to liveness detection as it is to biometric security [43] [114]. Whilst it follows most of the same factors, it also includes some additional concepts to be aware of. One of the most important, is due to the overall limited research conducted on the effects of medical conditions on liveness detection, such as medical noise data.

Universality identifies how widespread a sample is, therefore fingerprints and facial recognition are very universal is most people have both a face and fingerprints. Obviously there are a number of illnesses and conditions that can affect these samples as identified earlier. However the identification of medical implications will be technical in nature and no medical interpretation will be used, as [138] highlights, anyone that is not a professional in the medical field should refrain from any form of diagnosis or discussion.

Universality has a number of constraints to deal with, the main of which is how universal is a specific sample such as: irises do not have the ridges and valleys that are used within a fingerprint minutiae based detection algorithms [140] and therefore they cannot both be used using the same technique. However palm print techniques would be able to make some use of the fingerprint minutiae process. Therefore the degree of universality must be identified, some techniques are more universal than others as they can be effected by different liveness technique and can cross biometric styles.

This is where biometric and liveness universality differ as one is more susceptible to deviations. When considering biometric universality there are many effects that can cause problems, for example the effect of adermatoglyphia, or the lack of fingerprints is a very problematic, however the condition is very rare with approximately four extended families suffering from it worldwide. Therefore this is very unlikely to be an impactful factor within the security environment [33]. Whilst this is not a highly common issue, the concept is very pertinent, as most universality issues surrounding biometric devices can cause extreme problems for a security system, however the frequency and chance of them happening is very limited, therefore balancing out the danger.

In contrast, within liveness detection a simple pulse test can denote a living user, however as the average pulse is between 40-60 [117], and there are many factors that can cause deviations. These can be benign factors e.g. body position, temperature, exertion. To medical factors such as: heart arrhythmia; stroke; anxiety etc. Therefore it is reasonable to assume that there is a greater potential for inaccuracies, due to both the range and the prevalence of factors that can change the sample. Furthermore as liveness detection techniques are normally more medical in nature than biometric authentication or verification techniques they are more prone to illness/impairment etc. [166] [149].

A common feature within all liveness detection characteristics is the overall lack of generalisation, which is a stark contrast to the wide range of biometric fusion capabilities [114]. Multi-modal biometric systems can integrate different techniques together quite easily as they follow the same basic process, see Table 3.1, and there has been a lot of discussion regarding the fusion locations earlier within the research. Unfortunately due to the lack of a universal liveness detection techniques that encompasses different biometric security standards, this same degree of universality is difficult to attain. Therefore a combined universality categorisation for different techniques would allow a greater degree of liveness fusion and a more robust system. A good example of this is a fingerprint system that tests for salinity, which would not be suitable for an iris recognition system due to the lack of testing medium, which could cause issues to occur when creating a multi-model-based system. If the taxonomy was used, then the universality indicator would provide a usable comparison and would be able to identify similar levels of capability. For example the taxonomy would identify that fingerprint technique 1 has a 'level two' value for universality and iris technique 1 has a 'level 3' value for universality therefore enabling a statistical comparison of the techniques. This would make it easier to create more robust systems that can interface more thoroughly and provide a thorough approach to the development of multi-modal systems, as it would enable liveness detection techniques to be chosen that are of a similar security strength and universality.

This leads to the second characteristic of liveness universality: can the technique cover different biometric devices and is it found in different techniques naturally or are additional stages required to gather the samples. This leads to concepts such as liveness heterogeneity i.e. how universal is the technique, is it found in a large amount of samples/techniques, is it usable across multiple techniques, such as saline testing being viable within fingerprint, palm print, vein scanning etc., but not viable in iris recognition [121] [133].

Consequentially there are a number of factors that universality characteristics must consider, such as how easy is data access, how universal the technique is, as well as how universal the technique is when considering a multi-modal system, and/or fusion environment.

Throughout the development of this metric each section will be represented by a unique identifier. These identifiers are abstract in nature and have no relevance elsewhere, instead they are just a representation for the algorithmic purposes.The overall classifier is split into four different components that will provide the overall universality level.

Additional hardware

Is any additional hardware needed for this liveness technique. There may be no additional hardware needs, whilst other techniques may require many additions to the system. Therefore

the amount of additional hardware will need to be measured when considering the universality of a technique, the more hardware needed the less universal the technique is. Therefore the following measure will be taken (AH - Additional hardware):

1. Level 1 - 0 AH
2. Level 2 – 1 AH
3. Level 3 - 2 AH
4. Level 4 – 3 AH
5. Level 5 - $\geq$ 4 AH

This will then be assigned the letter ‘h’for the final algorithm.

Additional software

Whilst very similar to additional hardware, software demands its own category as it has some unique aspects that can affect it. What factors need to be included, are updates needed, additional installations etc. The more factors that are needed to gather the sample the less universal it will be. The ideal is that a sample can be gathered without any additions to the system. Due to the similarity between this and the additional hardware metric, the classification is very similar. Therefore the following measure will be taken (AS - Additional Software) (UP - Update):

1. Level 1 – 0 AS
2. Level 2 – 1 UP
3. Level 3 – 1 AS
4. Level 4 – 1 AS & 1 UP
5. Level 5 – $\geq$ 1 AS and $\geq$ 1 UP

This would then be provided the letter ‘s’.

Technique Heterogeneity

How many other biometric techniques does this one work with,. This is important as one of the current flaws within liveness detection is that there are a lots of techniques that are for very specific technique and very few that are generic. Therefore how easily can the proposed liveness technique be used across multiple types of device, if it all. This identification will provide relevant data for the development of multi-modal fusion systems.

1. Level 1 – Works with ≥ 6
2. Level 2 – Works with ≥ 5 and ≤ 6
3. Level 3 – Works with ≥ 2 and ≤ 4
4. Level 4 – Works with ≤ 2
5. Level 5 – Works with 0.

This will be given the letter ‘T ’.

Inheritance

The final metric is to identify if the sample being checked can provide liveness detection data during the initial security scan. Therefore can the liveness sample be gathered from the security sample, or can the liveness data be gathered from the sample, at a later time. The later will always cause more university problems as it will require the user to provide a second set of data, therefore causing problems relating to acceptance and device transparency. The more intrusive a sample is, then the less universal it is . Unlike the previous elements this metric component will only have two states:

1. Initial sample
2. Additional sample

They will be identified with the letter ‘I ’.

Metric

This metric culminates in Figure 4.1 which identifies the overall level of universality a liveness detection technique has, this equation will then be linked with the other equations to create an algorithm that can highlight the overall level of security, or suitability a technique has within a system.

$$U = \frac{h+s+T}{I}$$

Fig. 4.1 Universality equation

This first three values are added together, this represents that each of the values has the overall importance, whilst the final value (inheritance) can have a greater impact as it will denote how easily the sample is gathered. If the sample is part f the initial sample then there is no additional problems using the system, whilst if an additional sample is required it is much harder to gather as the universality of the technique must be reconsidered to identify were this sample is coming from. To test this equation the data in Table 4.3 will be considered:

ID	Technique	h	s	t	I
1	Hippus dilation test	1	2	3	1
2	Skin conductivity test	3	2	5	1
3	Eigenface analysis	1	5	3	1
4	3d face analysis	4	1	2	1

Table 4.3 Permanence testing characteristics

If Table 4.4 utilises an iris scanning device and hippus dilation liveness test shown in Table 4.3. There is no need for additional hardware (h = level 1), one software update is needed (s= level 2), the technique can be used across three different biometric samples (t = level 3), and the liveness sample is inherent and present during the initial security sample gathering process (I = 1), as seen in Table 4.4. The other calculations can be seen in Table 4.4.

Process	Algorithm
1.	$U = \frac{1+2+3}{1} = 6$
2.	$U = \frac{2+2+5}{1} = 10$
3.	$U = \frac{1+2+1}{2} = 2$
4.	$U = \frac{4+1+2}{2} = 3.5$

Table 4.4 Universality calculation

Previously this technique would have given a ordinal high/medium/low classification, but now with the interval values it can provide other relevant information such as technique one equals a '1' whilst technique two equals a '3'. This can now be compared to other techniques with empirical evidence denoting the inferiority and superiority of different techniques. When other techniques have had reliable data sorted throughout this system then it will not only

provide a more robust indicator of a techniques ability, but it will also allow a build-up of information regarding the comparative differences between levels, something that a ordinal technique cannot do., especially when included within a novel presentation environment such as a 3d time map.

4.3.2 Uniqueness

Uniqueness is oft identified as the defining factor of biometric security as a sample must be unique, to allow a correct match to be ascertained. Therefore the entire crux of biometric security lies within this area, is the sample unique and does it allow for correct authentication to occur [8] [7] [114].

Alternatively, as liveness detection is simply the detection of a living sample, and not the identification, authentication or verification of a user, the uniqueness of a sample is of no significance. As long as the sample has the relevant liveness characteristics that allow a correct identification of liveness to occur, then this is sufficient. If this data is to be used within specific environments, e.g. a medical based iHome which monitors pulse, then it would be assumed that the authentication or verification process would be completed within the security layer of the system whereas the liveness detection is merely a technique to prevent spoof data being submitted. This is the case because two unique users may have the exact same liveness detection sample, for example a pulse of 40-60 bpm, and therefore the difference between the individuals which would be of no consequence, as the sample is not attempting to authenticate. Therefore it is actually expected that some users will have identical liveness detection characteristics [115].

Subsequently these factors identify that uniqueness is not necessary for liveness detection, and it will not be used within this taxonomy. This is not done lightly, as the main thrust of this research is to create a more general system that can be used by other techniques in the future, such as coercion detection. This flexibility is paramount and the inclusion of non-essential categories will allow for potential security complications as more consideration must be identified when developing the overall security of a system. The removing of uniqueness as a category removes these potential complications, and provides a more streamlined model to apply to other areas.

4.3.3 Permanence

Whilst biometric security considers a stolen sample to be of the utmost importance, the same is not so for liveness detection as the emphasis is put on detection rather than verification of authentication. Due to the difficulty of stealing many liveness characteristics, e.g. you

cannot steal a brainwave, or pulse reading whilst denying access to the original user, it is not a factor that is a problem to consider. Instead the main factor to consider, for liveness, is what conditions can change a liveness sample to a sufficient level where it is no longer accepted during the matching process.

As liveness is heavily reliant on medical based data, the transient nature of the human body is a very important to consider. The potential for a biometric sample to be flawed is very high due to a range of factors such as illness, accidents, or even intentionally changing the sample. If a sample is easily altered then it is has a low permanence, this includes inclusive factors such as noise susceptibility or poor device/technique performance. One such example is vocal recognition, a voice sample can differ dramatically due to intra-user divergence, which can be caused by any number of reasons e.g. illness, dehydration, exertion etc. Alongside these biological concerns, there are other external factors to consider such as: the amount of noise within the sample collection area. In this scenario the permanence of vocal recognition is very low. Correspondingly the permanence of iris recognition would be very high, as there are comparatively few factors to cause change to sample. However even this techniques has some sample degradation problems such as cataracts and glaucoma as identified by [137].

Unfortunately whenever an emphasis is placed upon medical environments a plethora of consideration arise. These considerations are only limited by the number of impairments/illness/conditions that are relevant to the associated sample and technique.Many of the standard liveness detection tests rely on biological data that is mutable, and is very susceptible to medical divergence, such as a salinity test for fingerprint/palm techniques. For example, when considering salinity testing, the mutability of the sample depends completely on the level of sodium within the bloodstream, and correspondingly perspiration, as tested via the skin conductance response test as explained by [128]. Therefore the sample may differ dramatically over the course of authentication attempts due to any number of conditions such as diabetes and hyponatraemia [73]. Over a period of time a user's skin salinity levels may change dramatically due to illness, lifestyle, routine etc. making the permanence a vital factor to consider but one that is difficult to identify. Obviously some techniques are inherently less susceptible to these permanence deviations, but are often coupled with expensive extra hardware such as EEG machines.

Permanence does have a major disadvantage, which is that it is very difficult to develop an interval measurement metric, as the quantitative data range can contain a huge potential range of permanence altering factors such as medical, environmental, behavioural etc. This would have to be considered in much more detail, and the potential best route would be

to identify the 'genre' of permanence alterations, and classify it within each area. Whilst permanence is very similar for both liveness and security there are some difference.

The overall classifier is split into three different components that will provide the overall permanence level

Benign Susceptibility

The first factor is part of a pair benign and medical effects. Benign susceptibility covers factors that are not medical, therefore environmental, human error, etc. When considering benign techniques their temporary nature is a key factor, e.g. location temperature may change a skin conductivity test. However this may not, and often will not, occur at each sample gathering attempt. These aspects are often transient in nature and may occur regularly over period of time, or may be occasional occurrences e.g. someone may jog to work every day, providing a continuous occurrence, or someone may run to work because they are late for a meeting. Therefore the amount of factors that can effect the sample in this manner denotes its benign susceptibility. Therefore the following measure will be taken:

1. Level 1 - Susceptible to ≤ 1
2. Level 2 - Susceptible to ≥ 2 and ≤ 3
3. Level 3 - Susceptible to ≥ 4 and ≤ 5
4. Level 4 - Susceptible to ≥ 6 and ≤ 7
5. Level 5 - Susceptible to ≥ 8

They will be identified with the letter 'B_S '.

Medical Susceptibility

The second half of the pair follows a very similar style but the emphasis is on the medical effects that can negatively affect permanence. It could be argued that the medical effects are more important because they have the potential to be longer lasting, have greater impact etc. and whilst this is often true, the focus of this research is the quantity of factors that can impact a sample and not how much each medical factor will effect the sample. Therefore the following measure will be taken:

1. Level 1 - Susceptible to ≤ 1
2. Level 2 - Susceptible to ≥ 2 and ≤ 3

3. Level 3 - Susceptible to ≥ 4 and ≤ 5
4. Level 4 - Susceptible to ≥ 6 and ≤ 7
5. Level 5 - Susceptible to ≥ 8

They will be identified with the letter 'M_S '.

Profile Integration

The integration of profile gathering techniques can potentially alleviate some of the problems permanence experiences. Normally a sample range will be highlighted then depending on the sample comparison to this range a positive or negative response will be created. This range will normally be identified by taking an average of the relevant measures, such as according to the NHS [117], the average heart rate/pulse range is 40-60bpm for an average adult. However if there are users that have constant pulse samples that are above or below these figures then there will be constant failure to accept scenarios occurring. One way to deal with this would be to incorporate profiling using a autonomous profile architecture, which would allow the technique to dynamically adapt using context awareness techniques, depending on the user's profile. This would provide a number of benefits including a more heterogeneous system, there would be less problems dealing with users whose samples are outside the system averages, and it would provide a dynamic adaptive environment for both new and current users.

The profile generated about the user would involve a lot of potentially sensitive data and therefore would introduce a number additional considerations. First and foremost is the ethical considerations that revolves around the gathering of medical and personal data, secondly by enlarging the potential range of data to accommodate these average deviations will potentially open additional security threats when using spoof data, as well as a host of security based concerns regarding the safe storage of this data. All of these factors are completely relevant and important to consider. However within this research this is not were the focus will be placed.

Therefore the following measure will be taken:

1. Level 1 – No integration
2. Level 2 – Can store data
3. Level 3 – Some integration at technology or profile level
4. Level 4 – Full profiling system or autonomous control

5. Level 5 – Full integration providing dynamic profiles and adaptive security.

They will be identified with the letter ‘P ’.

Metric

This metric culminates in Figure 4.2 which identifies the overall level of permanence a liveness detection technique has.

$$P_m \frac{B_s + M_s}{P}$$

Fig. 4.2 Permanence equation

This equation takes the first two factors, adds them together, then divides it by the level of profile integration. This is because the initial characteristics have similar impacts, whilst he inclusion of a full profiling system will reduce the impact of permanence. This is why P is dividing the other factors. Therefore the higher the profile integration the lower the overall output of equation. If Table 4.5 is considered:

ID	Technique	Bs	Ms	P
1	Hippus dilation test	2	4	1
2	Skin conductivity test	3	2	5
3	Eigenface analysis	1	5	3
4	3d face analysis	4	1	2

Table 4.5 Permanence testing characteristics

Then the outputs in Table 4.6 will be calculated.

Process	Algorithm
1.	$P_m = \frac{2+4}{1} = 6]$
	$P_m 5 = P_m \rightarrow 5$
2.	$P_m = \frac{3+2}{5} = 1$
3.	$P_m = \frac{1+5}{3} = 2$
4.	$P_m = \frac{4+1}{2} = 2.5$

Table 4.6 Permanence algorithm examples

This identifies that the skin conductivity test is the technique that has the most permanence which means that there are fewer potential deviations that can affect skin conductivity

response tests. It is also possible to highlight that skin conductivity response tests are twice as permanent as eigenface analysis. The interval based analysis can provide a researcher with much more specific information.

4.3.4 Collectability

The primary contact a user has with a biometric device occurs when the sample is collected, and whilst the emphasis is often put on the technological considerations, such as the integration of hardware/software etc. into a system, this is not the total consideration within biometric based systems. Unlike traditional environments that depend on user mental involvement, i.e. the remembering of as pass-code, biometric environments demand that the user provides some active characteristics to denote proof, normally in the form of medical information. Therefore the user must provide the sample, which can be done with the user's knowledge, for example fingerprint scanning, or potentially without the user's knowledge, for example facial recognition. Biometric techniques however require specific hardware and/or software installations to correctly authenticate and verify users. Correspondingly liveness detection techniques, in some cases, require even more technology highlighting the issue of collectability.

One key factor that has made key based security system so successful is that security keys do not discriminate against users nor does it prevent anyone using it. Therefore having a very low susceptibility to noise data, a very high degree of ubiquity and accessibility, which can be seen by the length of time security keys have been in use and the general acceptance of it by the public. One other factor for security keys success is the ease of integration, as it needs no additional hardware and comparatively little additional software. However one factor to consider is that all forms of security demands the user to cooperate with the system, which makes the integration of transparency techniques even harder to accomplish. There are exceptions to this and a host of ethical and legal issues arise when data gathering occurs without permission. Despite this biometric security suffers from a number of other collectability issues such as a user's willingness to provide samples or a user's inability to provide samples.

The ease of sample collection depends completely on the user's interaction with the system, therefore allowing certain techniques to act more transparently than others. These techniques endeavour to minimise human device interaction, therefore providing a more transparent system. For example a fingerprint scanner can be comparatively transparent when compared to others techniques such as vein scans or retina/iris scanning, however it still is an invasive technique that relies upon the cooperation of the user. Therefore robust integration of the more transparent technique must be considered when developing the overall system.

Whilst there are techniques that are have the potential to be completely transparent they encounter a plethora of legal and ethical pitfalls surrounding their practice and implementation, as seen with the facial recognition techniques employed during the 2002 American Superbowl [36]. Therefore the ethical and legal factor must always be considered when including transparent biometric techniques into a system, however this is not something that will be focused on during this research. There are multiple considerations when attempting to make transparent techniques as [186] postulates. The main of these is to provide a lightweight client to the user, therefore reducing the amount of input needed, and making the process seem second nature, allowing the user to utilise the technique without having to actively consider it. Unfortunately it is quite difficult to create systems that are this transparency or invisible, as identified by [174] and this is made even more difficult within biometric environments, as most techniques are very intrusive in nature. When combined with the ethical and legal issues that occur when biometric system are made too transparent as identified by [84] alongside the issues surrounding liveness detection such as the deployment of additional hardware, software, the requesting of additional samples etc. therefore becoming less transparent and relying more on the user's cooperation.

When identifying collectability the main factor to consider is what kind of sample is being collected and how.

1. Is the sample an intrinsic part of an the initial security sample, such as a pulse from a fingerprint scanner, and if so is the hardware capable of gathering the data during the intimal security sample collection event or is a second scanning required.

2. Is the sample gathered innately from the security sample, such as within ECG techniques as the security sample innately denotes liveness [32].

3. Is the system utilising a completely different liveness detection characteristic, such as a skin conductivity technique used within an iris recognition biometric security system.

To measure collectability, an ordinal system of measurement would not be viable as it will providing no specific information on the degree of difference between techniques. Alternatively an interval measuring system will enable techniques to compared more easily and with a higher validity. For example it could be assumed that the addition of hardware would create the most secure sample collection environment, because the hardware can be custom fit to the system and therefore could be adapted to meet the exact requirements. However, this is not necessarily the most suitable techniques to follow. If a device is able to gather liveness data from the initial security sample then the efficiency of the system will improve as there will be no need for additional computation and stages. the problem here is

that this technique relies on the innate security sample collection and would only be possible with certain techniques due to the device being used. For example a fingerprint scanner that has inbuilt skin conductivity testers would be an efficient use of devices, however if the system demands an ECG, or hippus dilation test, then there is little choice but to include additional hardware. A secondary problem would also arise when dealing with multi-modal systems as the different techniques potentially could demand separate liveness techniques and, unless they are all based around the same sample, would require additional hardware to gather the samples. The main factors within collectability detail data acquisition and how easily this can occur, therefore the following factors have been included:

Innateness

The most efficient way to collect date is when the initial security sample is presented, therefore making collection at this stage innate to the original biometric sample.This would save on computation resources by providing data without the need for additional scans. For example the sample is automatically gathered during the enrolment process, e.g. if an ECG biometric technique is being used such as identified by [101], then the test itself denotes liveness, as a heart rate must be available for the authentication to occur. The problem with this technique is that it is highly dependent on the original sample and the liveness options associated, as some techniques are stand alone and do not integrate with others easily. This metric checks to see if the liveness sample can be automatically gathered from the initial security sample. If it cannot it then checks if the initial sample has the data it needs and if so a second scan can be taken of the original sample. Whilst this is acceptable it is not ideal as it negates some of the system transparency features of the environment, it opens up user acceptance constraints, and the addition of a stage can cause a security threat vector to become more prevalent. Therefore this metric checks to see if the liveness detection techniques is automatically gathered when testing for authentication and will use the follow classifications:

1. Level 1 – Automatic sample gathering

2. Level 2 – Not automatic but capable

3. Level 3 – No automatic sample gathering capabilities

This classification will be measured using the letter 'S_i '.

Time

When collecting samples time is one of the most important factors to take into account as the longer it takes to gather a sample the less efficient it is. this has been shown in the failed Blackstone fingerprint system [166]. This system forced the user to be still and took 6-8 seconds to authenticate, and was subsequently scrapped due to this poor performance. This shows that the longer it takes to gather a sample the less likely a user is to accept the technique and therefore the overall performance of the system will be impacted. Subsequently a time based measurement must be considered using the Blackstone experiment to identify the maximum time allowed to collect a sample. The following classifications will be used:

1. Level 1 – $\leq$ 1 second.
2. Level 2 – $\geq$ 1 & $\leq$ 1.5 seconds.
3. Level 3 – $\geq$ 1.5 & $\leq$ 2 seconds.
4. Level 4 – $\geq$ 2 & $\leq$ 2.5 seconds.
5. Level 5 – $\geq$ 2.5 second

This classification will be identified with the letter ‘T ’in the following equations.

Universality

Collectability is reliant on how easy a sample is to gather therefore it must consider what other techniques are applicable to a single liveness technique for example, does a fingerprint sample allow the use of a hippus dilation test. In this scenario the answer is no, however if a skin conductivity response test is then considered there are multiple techniques that can use this one test such as fingerprint, vein and palm scans. If the liveness technique is not universal there is an increased difficulty when gathering the liveness sample due to the multitude of liveness techniques in use. Therefore the more universal the liveness techniques the more easily they can be implemented into the overall system, as they can effect multiple devices. For example, fingerprint liveness samples can be gathered using perspiration, ridge frequencies etc., which can also be used across other techniques such as palm print and vein scans. Therefore the following classifications identifies liveness techniques that can also be used across other biometric sample collection methods:

1. Level 1 - $\geq$ 5 techniques
2. Level 2 - $\geq$ 3 and $\leq$ 5 techniques

3. Level 3 - ≥ 1 and ≤ 3 techniques
4. Level 4 - 1 technique
5. Level 5 - No techniques available

This classification will be identified with the letter 'U_t 'in the following equations.

Additional Technology

Like other factors, the addition of hardware and software to a system also adds a variety of threat vector. Ideally the sample collection should occur automatically when the sample if provided to the security process, however this often no the case and the system requires additional hardware and/or software to correctly gather data. For example, a sample my have the potential to be automatically gathered but requires additional hardware or software to make the technique work correctly. Therefore the final measure will identify what additional factors are needed to successfully gather the data(AH - Additional Hardware: AS - Additional software: UP - Updates):

1. Level 1 - ≥ 4 AH & ≥ 1 AS & ≥ 1 UP
2. Level 2 – 3 AH & 1 AS & 1 UP
3. Level 3 - 2 AH & 1 AS
4. Level 4 – 1 AH & 1 UP
5. Level 5 - 0 AH & 0 AS

This classification will be measured using the letter 'A_t 'in the following equations.

When these metrics are combined they will create the following equation which will calculate the overall level of collectability that will be comparable for different techniques due to the interval measuring system.

Metrics

This metric culminates in Table 4.7 which identifies the overall level of collectability a liveness detection technique has.

The initial values are added together as the have the same impact on the collectability of a sample, it is then multiplied by A_T as the impact initial techniques can have on the collectability of a sample is dramatic. The more additional technology required reduces the

$$C = \frac{S_i + T + U_t}{A_t}$$

Table 4.7 Collectability equation

ease of collection as more time is required along with more user cooperation. These factors require additional technologies to have a greater impact. To test this equation the data in Table 4.8 will be considered:

Technique	Si	T	Ut	At	Outputs
ECG	1	3	5	1	2.5
Skin conductivity Test	3	4	1	3	9.75
Temperature	1	2	2	3	4.5
Hippus Dilation	5	1	4	3	11.25

Table 4.8 Collectability Metrics with example data

Then the outputs in Table 4.9 will be calculated.

Example	Calculation
1	$C = \frac{1+3+5}{1} = 9$
2	$C = \frac{3+4+1}{3} = 3$
3	$C = \frac{1+2+2}{3} = 2$
4	$C = \frac{5+1+4}{3} = 3$

Table 4.9 Collectability Examples

As Table 4.8 and Table 4.9 shows this identifies the ECG technique as level 9 whilst hippus dilation test level is 2. This means that ECG is over three times less collectable than the hippus technique. This can then be used alongside other data that shows the ECG techniques are difficult to use due to the massive amount of hardware required to gather samples.

4.3.5 Performance

The performance of a system is often seen as the most important factor therefore the development of efficiency is vital to provide the best experience to the users, and provide the most useful data to other aspects of the system. Biometric performance is measured using metric such as the FAR/FRR/FTE etc. as discussed within Chapter Three. Unfortunately there are no

such metrics for liveness detection which is mainly due to the lack of generic standards. This leads to a fragmented measuring system that does not function cross-technique, therefore the inclusion of the interval measure will provide useful data that can be used comparatively to choose the best technique for an area. As [113] states there is a need to identify a universal performance indicator as the inclusion of any performance metric will offer a more thorough understanding of both the metric, and the techniques used. A side effect of this interval measure would be that it would allow a more suitable identification of multi-modal and fusion-able techniques to be chosen. For example, a skin conductivity technique may have a higher degree of performance when combined with fingerprint and palm print, instead of with iris recognition.

Luckily measuring performance can be, potentially, quite simple as there is a Boolean based result system, a person is either alive or not, therefore degrees of similarity and familiarity are irrelevant, instead as long as a user's sample is within a specific data range authentication can occur. Each technique has its own method of discerning this liveness which can work across multiple techniques or be specific to one and whilst there is a simple goal for liveness there are still accuracy goals to strive for. For example if a hippus dilation test can detect a live user nine out of ten times, then it is more secure than a skin conductivity test which can only achieve eight out of ten times [149]. Problems begin to occur when identifying the range in which acceptance is accepted as there are a host of noise factors that can cause normally static signals to fluctuate wildly, this will have an effect on liveness detection when considering spoofing and circumvention techniques. The following classifications will be used:

Technique Heterogeneity

How many other biometric techniques does this one work with. This is important as one of the current flaws within liveness detection is that there are a lots of techniques that are for very specific technique and very few that are generic. Therefore how easily can the proposed liveness technique be used across multiple types of device, if it all. This identification will provide relevant data for the development of multi-modal fusion systems.

1. Level 1 - Works with ≥ 6 techniques
2. Level 2 - Works with ≥ 5 and ≤ 6 techniques
3. Level 3 - Works with ≥ 2 and ≤ 4 techniques
4. Level 4 - Works with ≤ 2 techniques

5. Level 5 - Works with 0 techniques

This will be given the letter 'T_h '.

Accuracy

Accuracy is always a extremely important factor within any security environment. Within liveness detection, the accuracy of the technique can be affected by factors such as spoof data and noise. Therefore it is of vital importance that a viable accuracy measure is identified using the interval system. However as there are no general liveness measurements and for the sake of continuity, the same measures will be used from device accuracy which equate to the FAR, and FRR. False accept rate will denote the times that the system accepts a living sample that is not legitimately correct and the false reject rate identifies when a liveness sample is rejected when it should not be rejected because of a matching error. As there is no authentication or identification within liveness detection, nor is there any comparison, the need for additional measures is minimal, or example there is no need for FTE as no user enrolling will occur. To gather the raw data to measure the FAR and FRR, the EER (Equal Error Rate) is calculated. This is done by creating a ROC curve with the data and then finding the place that the FAR and FRR intersects is the ERR. This is the way accuracy will be measured and is shown in Figure 4.3.

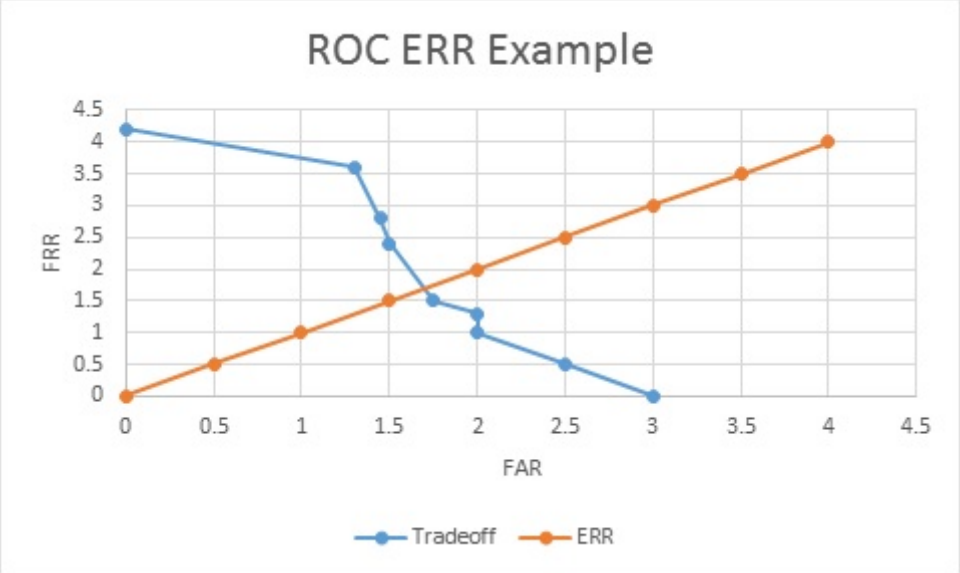


Fig. 4.3 Receiver operating characteristic (ROC) ERR

Therefore the following metric will contain the follow classifications:

1. Level 1 – ERR $\leq 1\%$

2. Level 2 - ERR $\geq 2\% \leq 3\%$
3. Level 3 - ERR $\geq 3\% \leq 4\%$
4. Level 4 - ERR $\geq 4\% \leq 5\%$
5. Level 5 - ERR $\geq 6\%$

Figure 4.3 shows the ERR being identified, this is done by plotting the FAR and FRR onto the graph which results in a progressive line. Then a standard line is plotted and were this standard line intersects the trade-off line, denotes the ERR. Therefore in Figure 4.3 the ERR would be 2% equating to a level 2 classification [163].

This classification will be identified by the letter 'A 'in the following equations.

Sample Spoofing

Ideally the accuracy fo the technique would be the only factor pertinent to sample collection. However this is not the case and there are other factors that can effect the performance of the systems, spoofing is one such factor. There are many threat vectors within biometric security however they do not crossover with liveness detection very well. However this is important to consider when dealing with liveness samples as the threat vectors associated with the techniques are extremely extensive as shown by [130]. Whilst most spoof samples are designed to beat the actual biometric security, and not the additional liveness detection, the spoofing of liveness samples is an area that is becoming focused on as liveness detection itself is becoming a standard addition to biometric devices [149]. Whilst liveness detection does not contain any authorisation or identification the spoofing of samples is still something to consider. The main problem being it does not matter who's sample is being tested, only what the sample is providing. Therefore the following metric measure how easily the sample is to spoof (ST - spoofing technique) e.g. audio pulse data:

1. Level 1 - 0 ST
2. Level 2 - 1 ST
3. Level 3 - ≤ 2 ST
4. Level 4 - ≥ 2 & ≤ 4 ST
5. Level 5 - ≥ 4 ST

This classification will be identified with the letter 'S_s 'in the following equations.

Additional Technology

Once again the need for additional technology can cause a performance degradation due to increased amount of system resources used. Therefore the following metric will conform to the earlier measures to minimise confusion (AH - Additional Hardware: AS - Additional software: UP - Updates):

1. Level 1 - $\geq$ 4 AH & $\geq$ 1 AS & $\geq$ 1 UP
2. Level 2 – 3 AH & 1 AS & 1 UP
3. Level 3 - 2 AH & 1 AS
4. Level 4 – 1 AH & 1 UP
5. Level 5 - 0 AH & 0 AS

This classification will be measured using the letter 'A_t '.

Time

As identified within previous classification the time it takes to check for a sample and process it, is a very important feature, as identified by failed projects that took too long to gather the data. Therefore the same time standards will be used:

1. Level 1 – $\geq$ 2.5 second
2. Level 2 – $\geq$ 2 & $\leq$ 2.5 seconds.
3. Level 3 – $\geq$ 1.5 & $\leq$ 2 seconds.
4. Level 4 – $\geq$ 1 & $\leq$ 1.5 seconds.
5. Level 5 – $\leq$ 1 second.

This measure will be identified by the letter 'T '.

$$P_f = \frac{T_h + A + S_S + A_t}{t}$$

Fig. 4.4 Performance equation

ID	Technique	T_h	A	S_s	A_t	T	Outputs
1	Hippus dilation test	5	1	2	2	2	5
2	Skin conductivity test	1	1	2	4	4	2
3	3d face analysis	2	2	5	5	5	2.8
4	ECG	2	1	2	5	3	3.3

Table 4.10 Performance testing characteristics

Metric

This metric culminates in Figure 4.4 which identifies the overall level of performance a liveness detection technique has.

Time is the most impactful value in this metric, that is why it is used to divide the rest of the calculation, the quicker the time the higher the value and therefore the lower the overall value. To test this equation the data in Table 4.10 will be considered:

Then the outputs in Table 4.11 will be calculated.

Table 4.11 shows that skin conductivity response test has the most performance and performs twice as well compared to hippus dilation test mainly due to the technique heterogeneity. It also shows the ECG and 3d face analysis tests are very similar which is surprising, as ECG is a a very new approach whilst 3d face analysis have been used for a number of years and are well researched within the area.

4.3.6 Acceptability

The acceptance of new technology is potentially an uphill struggle as it is often the case that technology, which is perfectly viable, will not be taken up due to an adverse public opinion [83]. Therefore the acceptability of any liveness detection technique must be relevant enough to allow a seamless integration, minimising any user acceptance problems. There are legitimate concerns that users may refuse to use, or be unhappy about using, certain techniques that they feel are invasive, unsafe, or insecure and whilst this can be alleviated somewhat with a good knowledge of the devices and a transparent identification of data flows there will always be some aversion to the technology. This is especially relevant when considering biometric security as they are a comparatively new security technique compared to pin numbers and passwords. Unfortunately the normal route the general public gets to hear

Test	Calculation
1.	$P_f = \frac{5+1+2+2}{2} = 5$
2.	$P_f = \frac{1+1+2+4}{4} = 2$
3.	$P_f = \frac{2+2+5+5}{5} = 3$
4.	$P_f = \frac{2+1+2+5}{3} = 3.3$

Table 4.11 Performance metric testing

about biometric devices is when they have been demonised in the media. This is done with a variety of erroneous results popularised within media which often are designed to provide entertainment without any thought regarding the realistic attributes of the technology.

Liveness detection has both advantages and disadvantages when considering acceptability issues and one of the primary disadvantages is that the name, liveness detection, denotes a medical emphasis which can potentially worry people who are unaware of its true purpose. To alleviate this issue, liveness detection must endeavour to be as transparent as possible, utilising as non invasive techniques as possible. By not providing the user with an outlet to notice the sample collection process can improve the overall acceptance and integration of the technology into daily routine. However there is an alternative view that this may produce ethical and legal ramifications this technique would encounter. Whilst to some users the transparency would make the technique more usable, others it would further alienate for fear of what the data is being gathered for and why. An excellent example of this is the XBOX Kinect, which was included into the games console the XBOX One, as it was identified that the Kinect would constantly capture data to improve the overall service, however due to colossal user backlash this factor was rescinded. Therefore if this technique transparency was to be used it would have to be made clear to the users that there will be some form of liveness detection gathering in progress, and that the data was only to be used for stated services. This is only relevant when considering inherit and innate characteristics, as anything that demands the user to provide additional data, or any technique that demands the user engage with additional hardware after the authentication process, will automatically provide the user with the knowledge of sample gathering.

It is almost impossible to identify people's opinions without some form of primary data collection, identifying what level of knowledge and acceptance a user base would have when dealing with a specific technique. This would allow a level of generalisation to be identified and therefore measured. Therefore to gather acceptance data a different technique must be considered when compared to the other metrics. Whilst the acceptance of a technique is a very subjective approach concerned mainly with qualitative data, it is a simple task to convert this into a numerical measure by utilising a simple ranking system such as the Likert system.

If an interval measuring system is to be adopted then the focus for acceptability would be to correspond the previous ordinal rankings into a numerical format.

1. Very satisfied 80%-100%
2. Satisfied 60%-79%
3. Ambivalent 40%-59%
4. Unsatisfied 20%-39%
5. Very unsatisfied 0%-19%

This simple numerical ranking system has been used to identify the different levels, one being the best whilst five being the worst. This will allow a two point low scale, a two point high scale and a single point medium scale. This would then be directly linked to a Likert scale based upon user answers, with the traditional levels denoting the different levels on the scale. This data can then be gathered and a medium found to denote a level of acceptance for a specific area. Whilst this has its flaws, without the integration of a wide scale testing environment it will provide a good indication of specific acceptances within the area.

Metric

This metric culminates in Table 4.12 which identifies the overall level of acceptability a liveness detection technique has.

$$P_f = \frac{V_s}{M_s} * 100$$

Table 4.12 Acceptance Equation

This identifies the overall level of acceptance a coercion detection techniques has. It is created by finding the percentage each rating has, which is then compared to the following list:

1. Level 1 80%-100%
2. Level 2 60%-79%
3. Level 3 40%-59%
4. Level 4 20%-39%

5. Level 5 0%-19%

To test this equation the data in Table 4.13 will be considered:

Technique	V. satisfied	Satisfied	Ambivalent	Dissatisfied	V. dissatisfied	Level
Fingerprint	25	1	1	1	2	1
Iris Recognition	1	25	2	1	1	2
Facial Recognition	1	1	25	2	2	3
Gait Recognition	1	1	2	25	1	4
ECG	2	1	1	1	25	1

Table 4.13 Acceptance example data (Max user – 30)

Each row will then be calculated to find the highest value.

TKT	IFA	SCP	PS	ECG
$A = \frac{V_s}{M_u} * 100 = 83\%$	$A = \frac{V_s}{M_u} * 100 = 3\%$	$A = \frac{V_s}{M_u} * 100 = 3\%$	$A = \frac{V_s}{M_u} * 100 = 3\%$	$A = \frac{V_s}{M_u} * 100 = 6\%$
$A = \frac{S}{M_u} * 100 = 3\%$	$A = \frac{S}{M_u} * 100 = 3\%$	$A = \frac{S}{M_u} * 100 = 83\%$	$A = \frac{S}{M_u} * 100 = 6\%$	$A = \frac{S}{M_u} * 100 = 3\%$
$A = \frac{A}{M_u} * 100 = 3\%$	$A = \frac{A}{M_u} * 100 = 83\%$	$A = \frac{A}{M_u} * 100 = 6\%$	$A = \frac{A}{M_u} * 100 = 3\%$	$A = \frac{A}{M_u} * 100 = 3\%$
$A = \frac{U_s}{M_u} * 100 = 3\%$	$A = \frac{U_s}{M_u} * 100 = 3\%$	$A = \frac{U_s}{M_u} * 100 = 6\%$	$A = \frac{U_s}{M_u} * 100 = 83\%$	$A = \frac{U_s}{M_u} * 100 = 3\%$
$A = \frac{V_u s}{M_u} * 100 = 6\%$	$A = \frac{V_u s}{M_u} * 100 = 3\%$	$A = \frac{V_u s}{M_u} * 100 = 3\%$	$A = \frac{V_u s}{M_u} * 100 = 3\%$	$A = \frac{V_u s}{M_u} * 100 = 83\%$

Table 4.14 Acceptance metric

This data identifies a sample group of 30 users who provide their views on the techniques mentioned. However to convert this to interval data is a little more complex. The data range for each level is denoted by 20% increments. Therefore the lowest level 5 would equate to 0-20% whereas the highest level 1 would be 80-100%. This enables statistical representation and analysis to occur, e.g. fingerprint techniques achieved a 83% rating which means it is level 1 , whereas ECG only achieves a 20% therefore giving it a Rank/Level 5 rating. This is done by finding out which areas the technique has the greatest value and then finding the corresponding level.

4.3.7 Circumvention

The ease of circumvention is one of the most highlighted points within biometric sample collection, the production of spoof biometric samples has been, and is, a constant problem to contend with. The ease of both collection and fabrication of spoof has become well known and it is of the utmost importance when considering the relevant biometric device

to include within a system. For example it is comparatively easy to access a fingerprint due to their ubiquitous nature, fingerprints can be lifted from many surfaces, including empty drink cartons, cups/glasses, windows, door-frames etc. When coupled with the plethora of information, easily accessible, regarding the creation of spoof fingerprints as identified by [171] [105] it shows a high degree of circumvention is possible which is also comparatively easy to replicate. This is not the case for all biometric techniques as some, such as ECG, are much harder to spoof however the spoof-ability of a technique is one of the primary concerns when choosing the correct technique to be implemented within a system.

Liveness detection circumvention is just as important, as no matter how robust and advanced a security system is, if a nefarious user is able to circumvent the defences the security is for nought. It does gain, and suffer, from the same attributes identified by the performance metric due to the lack of general liveness detection techniques. Each device has its own techniques and whilst some do share similarities they often use standalone liveness techniques therefore having a very low universality.

Sample gathering is important because if the user has to provide an additional sample to a separate piece of hardware then there is potential for a nefarious user to have prepared separate spoof sample to fool the system, therefore making the identification of threat vectors more difficult to discover. There are many representations of technique circumvention that covers biometric security such [171] [105] etc. and all of these techniques have common features. The susceptibility, to a lesser or greater extent, to spoof data, theft of sample, coercion and non-authenticated sample being proposed. It is these techniques that must be considered.

The main factor when considering circumvention is what manner the circumvention can encompass. When dealing with passwords and passkeys, there are many different threats, theft, carelessness, phishing, etc. [156]. Similarly liveness detection deals with different factors ranging from biological, chemical, behavioural etc. samples that provide the acceptance or denial therefore it is this sample that is the focus of circumvention, and the techniques that can measure it.

The development of spoof samples is a huge problem within biometric security as well as liveness detection, and liveness detection is one of the main methods to improve and defend biometric security from these spoof attacks. Techniques such as using HD video to spoof hippus dilation tests, and using a serum to spoof skin response tests, each technique shows the importance of good spoof data. Therefore attributes such as the permanence of a specific sample must be taken into account, as the less permanent a sample the greater range of circumvention can occur. A secondary characteristics denotes the applicable ranges of data a sample can be authenticated with, the bigger the range the more chance a nefarious user has

when trying to circumvent the system. For example using heart rate as a liveness monitor, heart rate averages between 40-60 beats per minute (NHS Choices 2013) however this can be changed dramatically depending on a huge variety of conditions and circumstances. Correspondingly when considering a more accurate technique such as the hamming distance between eye pixel depths within a 2d/3d environment as identified by [88]. This technique develops an average threshold then an average result is gathered, by measuring the amount of pixels between two frames, with any factor over the threshold then it is a living face, therefore. Due to the lack of high degree of permanence it proves to be a more secure technique then many others [19]. Therefore this aspect of the metric must identify what the comparative degree of precision is, by identifying how many permanence based factors can affect the sample, and how specific the sample can be. All of these factors will be considered when developing the overall metric:

Spoof gathering

The First factor to consider is how easy is it for a nefarious user to access a sample. Is the sample difficult to attain such as Purkinje image or is it easy such as fingerprints. The harder it is to gather this sample the harder it is to create a sample. Therefore the following classification will identify this value (SGT = Sample Gathering Techniques):

1. Level 1 – 0 SGT
2. Level 2 - $\leq$ 1 SGT
3. Level 3 - $\geq$ 2 $\leq$ 3 SGT
4. Level 4 - $\geq$ 4 $\leq$ 5 SGT
5. Level 5 - $\leq$ 6 SGT

This will be given the letter 'S_g'in the following equations.

Accuracy

The accuracy of the technique is key, as it allows the system to correctly identify a user, therefore it is a key area for an attack to occur. The higher the accuracy the harder the system is to breach.

1. Level 1 – ERR $\leq$ 1%
2. Level 2 - ERR $\geq$ 2% $\leq$ 3%

3. Level 3 - ERR $\geq 3\% \leq 4\%$
4. Level 4 - ERR $\geq 4\% \leq 5\%$
5. Level 5 - ERR $\geq 6\%$

This will be given the letter 'A'in the following equations.

Linked Sample

One unique issue that effects liveness detection, which is often seen as a positive factor within other classifications, is the potential for user sample similarity. Multiple users can have, and often are expected to have, the exact same liveness characteristic for example [3] identified a 90+% classification rate using a selection of industry fingerprint devices. However there is no indication to whom the sample belongs, and is therefore often deemed irrelevant. This focus is similarly identified in other works such as [44] and [88] focusing on only identifying the liveness sample and not the authentication. Whilst this is not always an issue, it must be identified and discussed. The traditional approach is that liveness detection occurs after a sample has been taken then the liveness detection process occurs. Therefore there is an innate degree of authentication that allows the system to trust the results. However if the liveness technique was removed from the biometric sample completely, e.g. facial liveness technique after finger and palm print authentication, then then there are potential security breach areas that occur. If the sample is from the same initial security sample then there are strict limitations placed on it which reduce the areas of threat. However if there are separate samples being gathered and they are not linked with the security sample, then there can be security compromises that occur. As there are only two constraints that are relevant here, authenticated and not authenticated, then there will only be two degrees of metric.

1. Level 1 – not linked to security sample
2. Level 2 - linked to security sample

This will be given the letter 'L_s 'in the following equations.

Metric

Therefore Table 4.15 identifies the overall equation for circumvention:

The effect a linked sample has on circumvention is very impactful, as it means it is harder for a nefarious user to gather spoof as the amount of time to gather one is minimised. This is why the metric is divided in this manner, the higher the value the greater the impact of the

$$C = \frac{S_g + A}{L_s}$$

Table 4.15 Circumvention equation

division meaning a less susceptibility to circumvention. Table 4.16 and Table 4.17 identifies the raw data values for the circumvention metric, including a number of factors gathered from the other metrics identified

ID	Technique	S_g	A	L_s	Outputs
1	Hippus dilation test	5	1	2	3
2	Skin conductivity test	1	1	1	2
3	Eigenface analysis	2	2	2	2
4	3d face analysis	4	3	2	3.5

Table 4.16 Circumvention testing characteristics

Then the outputs in Table 4.17 will be calculated.

Metric no.	Equation
1.	$C = \frac{5+1}{2} = 3$
2.	$C = \frac{1+1}{1} = 2$
3.	$C = \frac{2+2}{2} = 2$
4.	$C = \frac{4+3}{2} = 3.5$

Table 4.17 Circumvention metric testing

This shows that the techniques tested are very similar with skin conductivity response tests and eigenface analysis being equal whilst being almost twice as hard to circumvent as 3d face analysis.

4.4 Coercion Detection Development

The second area that the taxonomy must apply to is coercion detection. This is because coercion detection is the subsequent stage, after liveness detection, within biometric security and unlike liveness it has the capability to conform to standards from its integration, instead of applying standards after the development of techniques. However to completely identify all potential coercion techniques is impractical due to the lack of current research. Therefore

a set of original techniques will be considered within the scope of the taxonomy and they will conform to the categories within Table 4.19.

Name	Biometric	Technique	Merits	Flaws
Alternative Fingerprint Authentication	Fingerprint	Providing specific fingerprint which indicates coercion.	Easy to implement	Can be obvious
Blinking Detection	Iris/Retina	Excessive blinking Adaptive noise cancellation.	Difficult to detect	Noise cancellation problems
Vocal Error Insertion	Vocal	User provides specific keywords that will denote coercion.	Easy to implement Traditional	User memory. Can be obvious
Focusing Relevance	Iris/Retina	User focusing on specific area of the scanner, then using x/y coordinates and time, denotes coercion.	Suitable Easy to implement	Reliant on user ability
FACS Adaptation	Facial	Provides the relevant FACS value, which is predefined as coercion.	High accuracy. High obfuscation .	User knowledge is key

Table 4.19 IFA Samples

4.4.1 Voluntary Techniques

Tangible Key

Tangibility has been a very important factor when considering the medium of technology. This is because tangible products have constituted the main form of media consumption for human technology interaction. Due to this innate familiarity there is a predisposition to accept techniques incorporating tangible characteristics. Whilst tangible factors are often considered not to be as efficient as digital variations, there are precedents set that disprove

this. For example whilst most will agree that electronic storage, to store books, audio, video etc. is a more efficient technique due to improved sustainability, access and ease of storage etc. there are many users who still prefer tangible media for a number of reasons including, the tradition it represents, the value for money feeling, and the social convention inherent to the media. These attributes provide a positive inflection for the use of tangible keys. When considering security based tangibility one of the main techniques would be the 'panic button' which is ubiquitous throughout medical systems and environments. Therefore the inclusion of a panic key for coercion detection could provide a efficient technique for system developers. This does not mean only dedicated tangible keys could be used, as other techniques that utilised features such as mobile applications, keys, cards etc. can all be utilised to create a coercion identifier.

The main function of this technique would be to allow a user a way to voluntarily identify if they are being coerced when authenticating into a system. This technique relies on the user activating their 'key' which in turn will be picked up by the system and appropriate responses would be applied. This technique has some very positive features, the foremost being the certainty of coercion, as the user is the ultimate authority on their own degree of coercion and whilst it is possible that the user may mistake coercion from aggression, it is still easier to identify what is occurring and how to react. This removes the issues surrounding the detection of coercion, and potentially provides a major efficiency boost. One consideration would be to identify how this responsibility would effect the user. Due to the innate scalability and flexibility of this technique, it would be easily distributed throughout a system as the development of keys could range from simple click-able buttons, to advanced mobile applications, all of which could provide a user with the capabilities to deal with their specific scenarios. Therefore the implementation cost in both fiscal and system resources could be limited, creating a more efficient system. As with all user specific technologies careful consideration would have to be given regarding the relevant system policies and training requirements as correct use of the keys would be vital to system efficiency, as misuse would cause a host of problems within the system ranging from simple false accept and rejects values to intentional distributed denial of service attacks.

This style of technique has a lot of potential, however there are some problems that could be very difficult to overcome. If a user is being coerced then it is likely that the attacker will search throughout the pockets, confiscate phone etc. therefore depriving the user of the key, unless they have the presence of mind, and time, to locate and activate their device. Whilst there are methods to get around this, they are often convoluted or impractical in applications. For example a system that requires a periodic authentication to maintain the keys validity, the 'I'm alive' or heartbeat technique. Obviously this creates even more questions such

as what happens if the user is sleeping are there any environment noise constraints etc. Therefore whilst this technique has some potential, the negative aspects overshadow the positive substantially, therefore during the metric analysis phase, it is expected that this technique will score very poorly overall, with the only advantages being the ease of key development and the expected public acceptance.

Intended False Authentication

Intended False Authentication (IFA), whilst following the core concept of tangible key techniques, tries to negate some of the problems associate with the mutability of the tangible key itself. IFA affords the user the capability of voluntarily providing a false authentication sample to indicate coercion without the need for a tangible key [12]. If a fingerprint device is used as an example: during the authentication process the user would normally authenticate with the index finger, as identified by the system policy, however if the user is being coerced then they would instead use the fifth digit. This would be identified within the matcher module, and the coercion positive information would be sent to the decision maker, which would have the coercion and act accordingly. Whilst this fingerprint example is heavily simplified it highlight the main idea that could be applied to all other biometric devices. Some other examples are shown within Table 4.19.

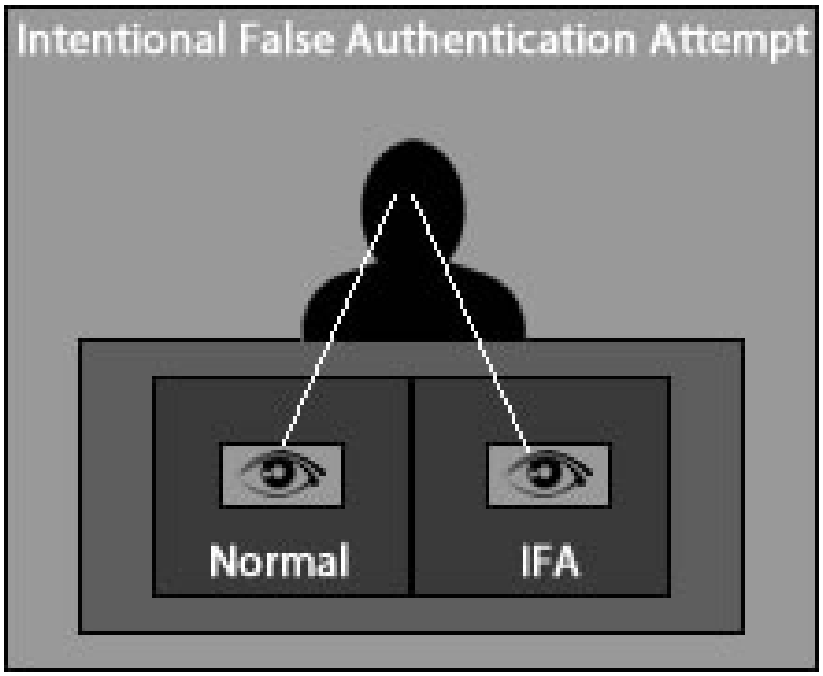


Fig. 4.5

Figure 4.5 shows an iris recognition system example, as it is quite simple to implement eye tracking for example using a Purine images. This differs somewhat from the previous fingerprint example, as eye location is not normally a biometric identification technique, instead it is counted as part of a sample that is used within noise cancellation algorithms. This process of Focusing Relevance would begin by dedicating a specific area for the user to look at. This would normally be located in a similar area to the iris recognition, therefore whilst iris recognition was occurring, the focusing on a specific zone would provide the system the evidence of coercion. This would be a separate process to the normal biometric authentication process and would have to be thoroughly integrated into the overall techniques, as failure to do so would disrupt the overall techniques and potentially cause problems for the user.

4.4.2 Involuntary

Skin conductivity peaks

Skin conductivity is a technique that has a high degree of universality and has been utilised in a number of liveness techniques for different devices. This technique uses the skin conductivity peaks during times of high emotion, and for use within coercion detection this would specifically be fear and similar negative emotions such as disgust which can provide similar results [68] [160]. One factor to remember: the view of affect testing is commonly misunderstood as emotion testing instead it is the physiological responses to emotions which are being measured, and it is these reactions that provide the potential coercion detection techniques [158] [160].

Skin conductivity response tests have some excellent attributes for example the response time of the test is normally very fast and there is a high degree of universality due to the plethora of liveness and coercion techniques that can utilise a SCR test. For example data gathered using fingerprint scanners, palm print, vein scanning etc. can be implemented alongside skin conductivity response tests. Additionally as this technique is as universal within liveness detection as coercion it can prompt fusion between the two areas which will culminate in a high standard of intra-technique universality [68]. Despite these positive factors one major disadvantage is SCR's susceptibility to noise, for example the effect the various cosmetics and skin care treatments may have on skin conductivity tests [157]. A secondary problem is that whilst this technique is very effective, it is also notoriously difficult to test, mainly due to the tricky situation of encouraging high emotions in users, dealing with the legal and ethical issues, etc. There have been examples of research were up to 10% of the participants refused to continue with the research, whilst it was in process due to the high

emotions being tested [68]. However despite this the technique has a lot of potential and can be included in a variety of different instances therefore showing that this universality is a major deciding factor. As [93] indicates, whilst SCR tests are one such technique to identify effect data, it is by no means the only one, however sufficient testing of any technique must be conducted before the final implementation of it is decided on.

Facial Micro-Movement

Within emotion detection one area is often discussed as potentially viable, as it allows for a diverse integration into technology, and has a robust research focus at its core. Facial micro-movement attempts to identify what muscle movements within the face can indicate an emotion using a predefined list of movements. If the stimuli for these muscle movements is a negative emotion then it can potential denote coercion. These movements are involuntary therefore ignoring a lot of the problems related to voluntary factors such as user error etc. However they do have to contend with a number of issues such as the reliance on affect detection, which as has already been mentioned, can be very imprecise.

There are a selection of techniques to facilitate Facial Micro-Movement (FMM), such as FACS (Facial Action Coding System) [54] and MAX (Maximally Discriminative Affect Coding System) [78]. The most viable of which is FACS as it is the most accurate [146]. This is due to the reduced amount of potential muscle identifiers that the MAX system uses, alongside the reduced accuracy [42]. Another issue is that the interpretation of emotion can differ between systems and different techniques sometimes cannot agree on [42] [146]. Therefore when considering FMM FACS will be used and whilst most of the research done into this area is purely within the emotion detection environment, there are computing precedents which, whilst not within the biometric area, still look at the relevance of FACS within computing [70] [100].

FACS uses a system of facial movements that are measured in Action Units (AUs) which are translated into emotions and are shown in Figure 4.6. These AUs are normally correlated by code, even though they can be correlated by observers as shown in [42]. However this technique has often been plagued with problems: as often two different programmers would not judge a sample emotion to be the same which would result in consistency issues. Therefore it must be identified at an early stage of develop what different AU combinations specifically mean then this metric would be used as a standard throughout all coercion techniques.

To gather this data FACS is made up of thirty AUs, and fourteen miscellaneous actions [54], however a fact that is often misunderstood is that there are two definitive techniques involved with the capture of this AU data and it is possible to obtain additive and non-additive

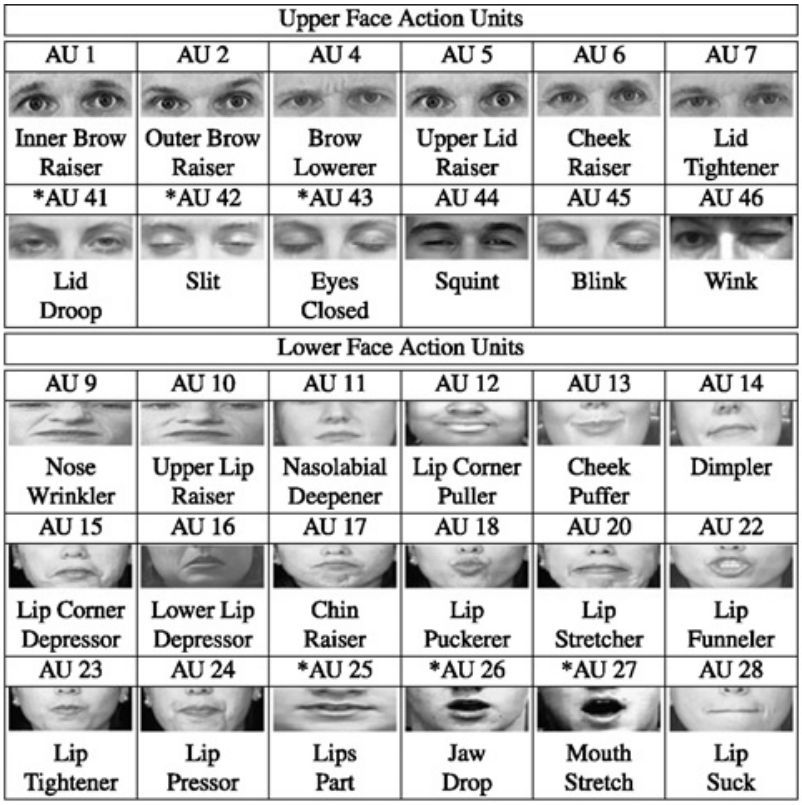


Fig. 4.6 AU Facial images [98]

samples, as identified by [164]. Additive techniques represent multiple AUs that are taken in conjunction, to form more elaborate overall results, whereas non additive techniques are always taken singularly. Correspondingly whilst additive techniques are more robust and effective due to the higher range of potential features, they have to content with a much wider selection of data collection problems and a more complex coding development [164].

Whilst the main advantages of FACS is the detection accuracy it provides, like many facial techniques it is prone to noise variations such as rigid head movement and face occlusion which is a common factor throughout facial biometrics [7]. Therefore an emphasis must be to minimise the effect of noise, as it can seriously hinder the relevant collection of data within this technique. One such technique is to use wide camera angles when capturing data, however this does increase the complexity as there are many variables for the technique to handle. This can lead to more problems throughout the overall technique such as potential fusion areas and applicability [146].

The accuracy of this technique must be also be considered. [146] indicates that FACS has achieved a reliable 90% success rate when dealing with two different variations of data. The first variation required the user to provide a predefined facial structure such as angry or happy. These results indicated that samples gathered from this collection style were often asymmetrical as the user's was trying to consciously achieve the required emotion and therefore the was some facial distortion due to the effort of controlling their facial muscles and consciously maintaining these expressions. Whilst there is potential for this area to be used within coercion detection, especially due to its voluntary nature, it is not practical due to the difficultly in replicating specific AU patterns.

The second technique utilises involuntary data which is much more symmetrical due to unconscious muscle movement alongside a quicker response. This technique can provide some very useful coercion detection data as the user can be examined by the system whilst they are authenticating and a selection of AUs can used to denoted coercion e.g. is the user authenticating sad, angry, frightened etc.

4.5 Coercion detection categorisation development

Throughout this research the intent has been to use the developed taxonomy to analyse liveness and coercion detection. Throughout the liveness taxonomy section each classification has been discussed and its merits identified, subsequently the next stage of the research is check for suitability with coercion detection. This will be done using a selection of novel coercion techniques that have been developed for this research. The following section takes the liveness detection taxonomy and adapts it for use within coercion detection, this novel

taxonomy will identify the salient points of a coercion technique by categorising the different factors inherent which will culminate in the development of an algorithmic output enabling a measure of suitability within specific environments.

However as some of the sections in this chapter are similar to the liveness standards no reiterate will occur, instead it will focus on the coercion aspects within the taxonomy as well as identifying that algorithmic representation therein.

4.5.1 Universality

Coercion detection requires a high degree of universality so that it can be relevant across numerous biometric devices. Therefore whilst it is not as vital as liveness detection it still has a robust part to play, influencing ease of integration, and heterogeneity of technique. Coercion detection techniques can provide a high or low degree of universality, and unlike liveness detection can contain techniques that have no direct correlation to the sample being provided. Tangible keys, pass-codes/words etc. are extremely universal as they can be used regardless of the individual sample gathering technique and have comparatively few medical factors that can prevent their use. As the speed of collection is one of the paramount factors the more universal a sample the better as there will be little time available for a user to provide the coercion data, and if the technique is not rapid and simple then coercion may not be identified alongside other authentication and detection problems occurring within the system.

Coercion detection potentially can be very universal due to the broad range of techniques that can be applied to biometric devices. For example a Focusing Relevance technique is applicable to any biometric that utilises a camera such as facial recognition, iris/retina recognition etc. Furthermore due to the comparative inexpensiveness of some additional hardware components such as web cameras etc., the ease in which different techniques can be implemented into other biometric or liveness systems is very high. For example including a Vocal Error Insertion technique would only require a microphone and the relevant software: comparatively easier to implement than an additional liveness detection standard. Therefore universality will be broken down into the following four areas:

Additional Hardware

Unlike liveness the main focus of coercion detection is to identify if the samples provided are from a coerced user subsequently there is an expectation that the sample provided will be from either the biometric security or liveness sample. If this is not the case there may be need for additional technology components. Whilst normally minor software adaptations should be

able to achieve the desired comparison sometimes this is not viable and additional hardware is needed. For example the alternative Fingerprint Authentication Techniques would require software adaptation to recognise that a pre-specified finger being presented will indicate coercion: whereas to detect liveness additional hardware would have be included for example a blood pressure monitor. Whilst this is not an immutable rule it is a generalisation therefore instead of including the liveness detection additional hardware metric the device adaptation metric will be used. This will include additions to software, as well as basic hardware and software adaptations amongst the lower levels of the metric. These will have the least impact on the system, due to the ease of inclusion and development. The later levels of the metric would focus on full software development, and expansive additional hardware integration, new medical data collectors etc. The following classifications will be used (Addition Hardware - AH, Additional Software - AS, Updates 0- UP):

1. Level 1 - 0 AH
2. Level 2 – 1 AH
3. Level 3 - 2 AH
4. Level 4 – 3 AH
5. Level 5 - $\geq$ 4 AH

This will then be assigned the letter ‘h’for the final algorithm.

Additional software

1. Level 1 – 0 AS
2. Level 2 – 1 UP
3. Level 3 – 1 AS
4. Level 4 – 1 AS & 1 UP
5. Level 5 – $\geq$ 1 AS and $\geq$ 1 UP

This would then be provided the letter ‘s’.

Heterogeneity

The degree to which the techniques can be implemented across different biometric and liveness styles can be identified with this classification. Whilst it would seem logical to ignore techniques that did not provide a higher degree of universality, this is not the case as there can be more specific techniques developed that work exceedingly well within a specific boundary. Nor does a high degree of universality mean a very effective overall technique. For example Tangible Key Techniques are very universal as they do not rely on the sample data, however they have a variety of other problems such as the potential ease of circumvention and the comparative collectability issues inherent. The following classifications will be used:

1. Level 1 - Works with ≥ 6 techniques
2. Level 2 - Works with ≥ 5 and ≤ 6 techniques
3. Level 3 - Works with ≥ 2 and ≤ 4 techniques
4. Level 4 - Works with ≤ 2 techniques
5. Level 5 - Works with 0 techniques

This classification will be identified with the letter 'H_t'in the following equations.

Inheritance

Like liveness detection a decision must be made to identify if the sample being used for biometric/liveness can be used for coercion detection. The potential separation of samples can be problematic however it can also boost security, as the data can be specifically adapted to be efficient, for example within Tangible Key techniques. However the reverse is equally as valid as when the sample is integrated into other techniques it can potentially be more efficient, and it does not create any new security concerns like a fully separated technique would do. Unlike liveness detection coercion necessitates three different levels within this metric, as they represent the three states a technique can exist in: completely separated(Tangible Key); Direct Association (skin conductivity tests from fingerprint biometrics); Indirect Association whilst not directly related to the sample can be gathered using the same kind of devices for example FACS Adaptation that is used with a iris/retina scanner.

1. Level 1 - Complete association.
2. Level 2 - Direct association.
3. Level 3 - Indirect association.

This classification will be identified with the letter 'I_h'in the following equations.

Metric

This metric then culminates in Table 4.20 which identifies the overall level of universality a coercion detection techniques has.

$$U = \frac{h + s + H_t}{I_h}$$

Table 4.20 Coercion universality equation

The level of inheritance is the most impactful factor in coercion universality, this is because of the importance surrounding the gathering o the sample. The more associated the sample the more efficient the system will be. Therefore this value will divide the sum of the other values. The closer the sample association the lower the overall level of universality. To test this equation the data in Table 4.21 will be considered:

ID	Technique	h	s	H_t	I
1	FACS	1	2	3	1
2	Skin conductivity peaks	3	2	5	1
3	TKT	1	5	3	1
4	IFA	4	1	2	1

Table 4.21 Coercion universality testing characteristics

If Table 4.4 utilises an iris scanning device and a FACS coercion test, as shown in Table 4.21, is considered: there is no need for additional hardware (h = level 1), one software update is needed (s= level 2), the technique can be used across three different biometric samples (h_t = level 3), and the sample is inherent and present during the initial security sample or liveness gathering stage (I_h = 1), as seen in Table 4.4. The other calculations can be seen in Table 4.4.

Process	Algorithm
1.	$U = \frac{1+2+3}{2} = 3$
2.	$U = \frac{2+2+5}{2} = 4$
3.	$U = \frac{1+2+1}{3} = 1$
4.	$U = \frac{4+1+2}{1} = 8$

Table 4.22 Coercion universality calculation

4.5.2 Uniqueness

Coercion is very similar to liveness when considering uniqueness as there are no requirements for identification, authentication or verification. Instead all that is necessary is that it can be identified if a user is being coerced when they authenticate into the system. As mentioned earlier: uniqueness in biometric security is of paramount importance however due to the sample confirmation nature of both coercion and liveness detection it will not be a factor that has any relevance.

4.5.3 Permanence

Permanence denotes the potential degree of changeability a sample contains due to any number of reasons such as medical variations, external noise, interference etc. One such factor being the deviation of a sample over time which can be related to medical problems, ageing etc. Coercion is a factor that is not prone to change due to the emotion based physiological data being gathered. Whilst in some circumstances the reaction to fear can dramatically change over time [18], it is unlikely to occur within coercion sample gathering. Whilst there is minimal chance that this will occur the potential is still there and therefore it must be considered to maintain the flexibility of the taxonomy. Whilst it might be unlikely to happen if it was to occur then the effect of this permanence mutability could dramatically impact a system. For example: considering a FACs based system it would not impact the coercion sample gathering if the user developed cataracts, which could interfere with both biometric and liveness detection. Only medical issues that seriously prevented an accurate identification of coercion would be an issue, for example a facial injury, from whatever source, that prevented or altered the FACS AU classification. Therefore the following classifications will be used:

Benign Susceptibility

Benign susceptibility focuses on non medical factors that can effect coercion sample gathering. Examples such as how long passwords can exists within a system before it needs to be changed etc. It also considers the robustness of tangible keys, how easily are they damaged, what is their normal lifespan, what effect would a faulty device have etc. Therefore the following five point scale identifies the degree in which the technique is susceptible to benign factors. Examples could include: tangible key degradation, key rotation, pass-code reliability etc. The following classifications will be used:

1. Level 1 - Susceptible to ≤ 1

2. Level 2 - Susceptible to ≥ 2 and ≤ 3
3. Level 3 - Susceptible to ≥ 4 and ≤ 5
4. Level 4 - Susceptible to ≥ 6 and ≤ 7
5. Level 5 - Susceptible to ≥ 8

This will be given the letter 'B_s'in the following equations.

Medical Effects

Medical implications cover the majority of permanence altering factors. Medical factors can effect the physiological characteristics that most coercion detection techniques are based on, concepts such as skin conductivity, heart rate variance, FACS etc. Therefore emotion detection can be negatively affected due to facial injuries caused by accident or illness. This will measure how susceptible a particular technique is to medical deviations, the more medical factors that effect the sample the higher the level therefore reducing the overall degree of permanence. The following classifications will be used:

1. Level 1 - Susceptible to ≤ 1
2. Level 2 - Susceptible to ≥ 2 and ≤ 3
3. Level 3 - Susceptible to ≥ 4 and ≤ 5
4. Level 4 - Susceptible to ≥ 6 and ≤ 7
5. Level 5 - Susceptible to ≥ 8

This will be given the letter 'M_s 'in the following equations.

Profile integration

This metric follows the same process as the one found in section 4.3.3 detailing the potential for dynamic profile creation allowing sample matching to take into account medical and benign deviations over time. Potential creating a more secure system due to the dynamic adaptation of individual users needs and reducing efficiency factors such as FRR and FAR.

1. Level 1 – No integration
2. Level 2 – Can store data

3. Level 3 – Some integration at technology or profile level

4. Level 4 – Full profiling system or autonomous control

5. Level 5 – Full integration providing dynamic profiles and adaptive security.

This will be given the letter ‘p ’in the following equations.

Metric

This metric culminates in Table 4.23 which identifies the overall level of permanence a coercion detection techniques has.

$$P_m = \frac{B_s + M_s}{P}$$

Table 4.23 Coercion Permanence Metric

Like liveness the integration of the results into a profiling system will improve the overall security within the technique. The greater the integration the high the division will be, therefore representing a more efficient system. This has the most impact on permanence within this algorithm. To test this equation the data in Table 4.24 will be considered:

ID	Technique	B_s	M_s	p
1	FACS	1	2	3
2	Skin conductivity peaks	3	2	5
3	IFA	1	5	3
4	TKT	2	1	3

Table 4.24 Coercion permanence testing characteristics

If the tangible key technique (TKT), from Table 4.24 is considered: it is susceptible to 2 benign techniques (B_s = level 2), 0 medical techniques (M_s = Level 1) and has some integration at technology level (p = Level 3). This would provide the overall output of 1, showing it to be a technique that is not prone to large permanence altering characteristics. The rest of the calculations can be seen in Table 4.25.

Process	Algorithm
1.	$U = \frac{1+2}{3} = 1$
2.	$U = \frac{3+2}{5} = 1$
3.	$U = \frac{1+5}{3} = 2$
4.	$U = \frac{2+1}{3} = 1$

Table 4.25 Coercion permanence calculation

4.5.4 Collectability

Regardless of the technical sophistication or relevance of a technique if it is difficult to obtain, then the impact of it is dramatically reduced. This is a concern regardless of technique and therefore it is unsurprising that the collectability of a sample has been the focus of much research throughout biometrics. Another reason for this focus is that it has received the most most media centric aspect of biometric security. The lifting of biometric samples to use for spoof development is a common issue and one of the main worries users have. This is because it is comparatively easy to gather biometric samples without the user's knowledge, for example lifting fingerprints from discarded cups, doors, other media etc., using HD images for facial and iris scanning etc. [171] [105].

Coercion is different because a number of the techniques do not use the same data that the security sample does. For example tangible key techniques do not suffer from spoof data development. Instead the user's key will be used by force, or a different key will be used. However some techniques still produce sample behaviours that consider: where is the sample gathered; is a second scanning needed; is it inherent to the original sample; is the sample automatically gathered as part of the previous scanning process; what additional characteristics are required to gather the sample data.

As with liveness samples the form the coercion sample takes is important to consider e.g. is it voluntary, involuntary, or environmental. Each generates different attributes depending on the situation for example: additional scans are more likely to be viable when using a voluntary samples, as the user is already prepared to provide additional information. Involuntary samples include the traditional sample collection techniques whereas environmental relies on non user centric factors for example cameras detecting user proximity, microphones detecting keywords etc. The following classifications will be used:

Innateness

As already identified the location of sample collection is a very important aspect to consider, the most efficient techniques will be gathered from the sample during the previous processes

(liveness etc.), however this is often not possible. Therefore this metric will identify where the sample will be collected from, the automatic collection of the sample during a previous scan is the most efficient as it does not required additional scanning, therefore cutting down on time required to authenticate the user.

1. Level 1 – Automatic sample gathering
2. Level 2 – Not automatic but capable
3. Level 3 – No automatic sample gathering capabilities

This will be given the letter 'I 'in the following equations.

Additional Technology

Like other factors the additional of hardware and software to a system also adds a variety of threat vector. Ideally the sample collection should occur automatically when the sample if provided to the security process, however this often not the case and the system requires additional hardware and/or software to correctly gather data. For example, a sample my have the potential to be automatically gathered but requires additional hardware or software to make the technique work correctly. Therefore the final measure will identify what additional factors are needed to successfully gather the data(AH - Additional Hardware: AS - Additional software: UP - Updates):

1. Level 1 - $\geq$ 4 AH & $\geq$ 1 AS & $\geq$ 1 UP
2. Level 2 – 3 AH & 1 AS & 1 UP
3. Level 3 - 2 AH & 1 AS
4. Level 4 – 1 AH & 1 UP
5. Level 5 - 0 AH & 0 AS

This classification will be measured using the letter 'A_t 'in the following equations.

Time

The necessity of quick processing is a standard feature throughout computing, there are many iterations of devices that have potential but are not quick enough to be viable, for example the PokerMetrics research developed by [158], which provide interesting results regarding

the physiological factors that could be used for coercion as it details results from high stress environments, however as the time it takes to gather data is approximately fifteen minutes it is not viable. Therefore the following classifications will be considered:

1. Level 1 – $\geq$ 2.5 second
2. Level 2 – $\geq$ 2 & $\leq$ 2.5 seconds.
3. Level 3 – $\geq$ 1.5 & $\leq$ 2 seconds.
4. Level 4 – $\geq$ 1 & $\leq$ 1.5 seconds.
5. Level 5 – $\leq$ 1 second.

This will be given the letter 'T 'in the following equations.

Metric

This metric culminates in Table 4.26 which identifies the overall level of collectability a coercion detection techniques has.

$$C = \frac{I + A_t}{T}$$

Table 4.26 Coercion collectability metric

Once again time is the main factor when considered collectability, the speed in which a sample is collected improves the overall collectability dramatically. Therefore this value is used to dived the rest of the values. The quicker the collection the greater the divisor will be. To test this equation the data in Table 4.27 will be considered:

ID	Technique	I	A_t	t
1	FACS	1	2	3
2	Skin conductivity peaks	3	2	5
3	IFA	1	5	3
4	TKT	2	1	3

Table 4.27 Coercion collectability testing characteristics

If a FACS coercion technique is considered alongside a facial scanning device, then Table 4.27 indicates that there is automatic sample gathering (I = level 1), three hardware.

two software and one update is needed (A_T= level 2), and the sample acquisition takes 1.6 seconds (T = level 3). This then shows that the output will equal 1, therefore achieving the highest level of collectability. Mainly due to the ease of gathering facial data cues when using facial recognition and the sample inheritance. The rest of the calculations can be seen in Table 4.28.

Process	Algorithm
1.	$U = \frac{1+2}{3} = 1$
2.	$U = \frac{3+2}{1} = 5$
3.	$U = \frac{2+5}{2} = 4.5$
4.	$U = \frac{1+1}{1} = 1$

Table 4.28 Coercion collectability calculation

4.5.5 Performance

Due to the limited availability of coercion detection data, few metrics of performance are available. Therefore the use of generic biometric measurements will improve heterogeneity and promote a comparative relationship with liveness detection. Therefore measurement such as FAR/FRR/FTE of will be used. Specifically the FAR will be used to identify when a coercion detection technique is unsuccessful therefore allowing a false user to be accepted as a non-coerced user, whilst coercion is occurring. This differs from the normal meaning of FAR which is when spoof data is being used to bypass the system. Within coercion detection the spoof data is not as important instead the actual acceptance is the main factor to consider. The FAR would mean that the matcher has not been able to detect coercion when there is data to indicate that this is occurring. The FRR would consist of users that the system believes are being coerced but are not, therefore stopping legitimate users into the system. This figure can be derived from a number of factors ranging from technical failures, such as skin conductivity response showing abnormal readings; noise fluctuation problems such as users exercise habits causing a FACS to identify them as being coerced.

Unlike other biometric system there are a number of techniques that really on user cooperation subsequently human error becomes a greater problem. For example a human may accidentally press their tangible key therefore providing evidence of coercion; the environmental techniques may detect to heat signatures close to each other, therefore constituting coercion, when it is simply people talking, or standing under an umbrella etc. So there are a number of unique concepts to consider within coercion detection. Therefore the following classifications will be considered:

Technique Heterogeneity

How many other coercion techniques does this one work with and how easily can the proposed coercion technique be used across multiple types of device, if it all. This identification will provide relevant data for the development of multi-modal fusion systems. Especially when considering coercion and liveness fusion.

1. Level 1 - Works with ≥ 6 techniques
2. Level 2 - Works with ≥ 5 and ≤ 6 techniques
3. Level 3 - Works with ≥ 2 and ≤ 4 techniques
4. Level 4 - Works with ≤ 2 techniques
5. Level 5 - Works with 0. techniques

This will be given the letter 'T_h '.

Accuracy

As mentioned the main focus of the accuracy classification is to utilise the main forms of biometric measurement, FAR and FRR. Whilst there are a much larger selection of classifications available, these are the main two which can provide relevant data for coercion detection. As future research continues the impact of other measurements such as FTE and FTA can be considered for integration. The use of a ROC diagram is the standard within biometric security as seen in Figure 4.3.

1. Level 1 – ERR $\leq 1\%$
2. Level 2 - ERR $\geq 2\% \leq 3\%$
3. Level 3 - ERR $\geq 3\% \leq 4\%$
4. Level 4 - ERR $\geq 4\% \leq 5\%$
5. Level 5 - ERR $\geq 6\%$

This will be given the letter 'A 'in the following equations.

Time

Time is always an issue to consider, as it can have a huge impact on the performance of coercion technique. This metric only considers the time taken to produce a positive or negative match for coercion and not liveness or security. Therefore the follow metric will be used:

1. Level 1 – $\geq$ 2.5 second
2. Level 2 – $\geq$ 2 & $\leq$ 2.5 seconds.
3. Level 3 – $\geq$ 1.5 & $\leq$ 2 seconds.
4. Level 4 – $\geq$ 1 & $\leq$ 1.5 seconds.
5. Level 5 – $\leq$ 1 second.

This will be given the letter 'T'in the following equations.

Additional Technology

Coercion detection differs from other biometric systems as there can be a larger selection of additional hardware or software due to the sheer variety of techniques available. Whilst it is similar to liveness it also includes a variety of individual features. The inclusion of techniques that relay on tangible devices, like panic keys or smart device proximity sensor, as well as a host of environmental additional such as cameras, microphones, proximity sensors etc. Therefore the potential of additional technology, regardless of its location, will increase. Subsequently the amount of resources the system needs to provide may fluctuate which can reduce the overall performance, as well as creating new threat vectors. Therefore techniques that utilise innate samples should, theoretically, perform better. Therefore the following classifications will be considered (AH - Additional Hardware: AS - Additional software: UP - Updates):

1. Level 1 - 0 AH & 0 AS
2. Level 2 – 1 AH & 1 UP
3. Level 3 - 2 AH & 1 AS
4. Level 4 – 3 AH & 1 AS & 1 UP
5. Level 5 - $\geq$ 4 AH & $\geq$ 1 AS & $\geq$ 1 UP

This will be given the letter 'A_t'in the following equations.

Benign Susceptibility

Benign susceptibility identifies noise deviations that are environment based, for example excess light, noise, temperature etc.

1. Level 1 - Susceptible to ≤ 1
2. Level 2 - Susceptible to ≥ 2 and ≤ 3
3. Level 3 - Susceptible to ≥ 4 and ≤ 5
4. Level 4 - Susceptible to ≥ 6 and ≤ 7
5. Level 5 - Susceptible to ≥ 8

They will be identified with the letter 'B_S '.

Medical Susceptibility

Medical susceptibility highlights medical noise deviations that can adversely effect coercion samples such hyperonatraemia for skin conductivity response tests.

1. Level 1 - Susceptible to ≤ 1
2. Level 2 - Susceptible to ≥ 2 and ≤ 3
3. Level 3 - Susceptible to ≥ 4 and ≤ 5
4. Level 4 - Susceptible to ≥ 6 and ≤ 7
5. Level 5 - Susceptible to ≥ 8

They will be identified with the letter 'M_S '.

Subtlety

The final aspect of this metric is a rather abstract one. How easy is it or the user to provide a coercion sample without the attacker being aware of the sample provision. This degree of subtlety is necessary to understand because if an attacker is aware of the user alerting the system to the coercion, then it could cause a host of problems for both the user and system alike. Different techniques have different degrees of subtlety, and the effects of subtlety can be irrelevant depending on the technique in question. For example any involuntary techniques would have a very low subtlety rating as neither the user or attack know that it is occurring,

$$P_f = \frac{T_h + A + A_t + B_s + M_s + S}{t}$$

Table 4.29 Coercion Performance Equation

whilst other techniques such as IFA would require the user to provide the data without an attacker knowing.

Therefore the following metric will identify the level of subtlety with involuntary techniques being the most subtle as the attacker has no way of knowing if there is a technique being implemented unless they have detailed knowledge of the system.

1. Level 1 - Both user and Attacker Unaware
2. Level 2 - User aware but sample automatically taken
3. Level 3 - User input, easy to hide – eye movement etc.
4. Level 4 - User input, hard to hide
5. Level 5 - System demands specific characteristics that differ from normal authentication techniques.

This will be given the letter 'S'in the following equations.

Metric

Time is the most impactful value in this metric, that is why it is used to divide the rest of the calculation, the quicker the time the higher the value and therefore the lower the overall value. This metric culminates in Table 4.26 which identifies the overall level of performance a coercion detection techniques has.

To test this equation the data in Table 4.30 will be considered:

ID	Technique	T_h	A	A_t	B_S	M_s	S	T
1	FACS	4	2	2	3	4	2	2
2	Skin conductivity peaks	3	1	2	2	5	3	4
3	IFA	3	2	3	2	2	5	1
4	TKT	5	2	3	3	4	2	3

Table 4.30 Coercion performance testing characteristics

Considering skin conductivity tests: it is viable across three techniques (T_H = level 3) , it has an ERR of 1.2% (A = level 2), it gathers data in 1.4 seconds (T = level 4), it requires one

hardware addition (A_T = level 2), it is susceptible to 2 benign and medical effects (M_s = level 2 and B_s = level 2), and it has a subtly level of 2 as the user is aware of the test but does not have to do anything to facilitate it (s = level 2). The equates to a performance rating of 4, which is the best tested being four times as efficient as IFA techniques, mainly due to the speed it captures data. The rest of the calculations can be seen in Table 4.31.

Process	Algorithm
1.	$U = \frac{4+2+2+3+4+2}{2} = 8.5$
2.	$U = \frac{3+1+2+2+5+3}{4} = 4$
3.	$U = \frac{3+2+3+2+2+5}{1} = 17$
4.	$U = \frac{5+2+3+3+4+2}{3} = 6.3$

Table 4.31 Coercion performance calculation

4.5.6 Acceptability

The public connotation surrounding both liveness detection and coercion detection is based around the understanding of the name. Popular media have portrayed biometric systems in a poor light and as easily broken oft with shocking techniques such as using cadavers or users being coerced into authentication. As coercion deals with a state that deals with unpleasant events, being coerced, many users may be reluctant to provide data. When coupled with the normal data range for coercion, i.e. data related to fear, disgust and other potential negative emotions, this factor is exasperated. Other acceptance factors can include: how users deal with this data being gathered; what are the ethical ramifications of having this data; what are the legal issues herein. All of these factors can reduce the public readiness to accepts coercion detection techniques. This can be tempered by identifying, to the user, that all possible eventualities are being catered for. One of the positive aspects of coercion detection is that there are acceptable techniques such as tangible keys and this could be used as a gateway to other techniques, providing the users with greater exposer to other lesser known techniques

Metric

This metric culminates in Table 4.32:

This identifies the overall level of acceptance a coercion detection techniques has. It is created by finding the percentage each rating has, which is then compared to the following list:

$$P_f = \frac{V_s}{M_s} * 100$$

Table 4.32 Coercion Performance Equation

This will then be compared to the following list to denote a rating as shown below:

1. Level 1 80%-100%
2. Level 2 60%-79%
3. Level 3 40%-59%
4. Level 4 20%-39%
5. Level 5 0%-19%

To test this equation the data in Table 4.30 will be considered:

Technique	V. satisfied	Satisfied	Ambivalent	Dissatisfied	V. dissatisfied	Level
Tangible Key	25	1	1	1	2	1
Intended False Authentication	1	25	2	1	1	2
Skin Conductivity Peaks	1	1	25	2	1	3
Proximity Sensor	1	1	2	25	1	4
ECG	2	1	1	1	25	5

Table 4.34 Acceptance example data (max user – 30

Each row will then be calculated to find the highest value:

TKT	IFA	SCP	PS	ECG
$A = \frac{V_s}{M_u} * 100 = 83\%$	$A = \frac{V_s}{M_u} * 100 = 3\%$	$A = \frac{V_s}{M_u} * 100 = 3\%$	$A = \frac{V_s}{M_u} * 100 = 3\%$	$A = \frac{V_s}{M_u} * 100 = 6\%$
$A = \frac{S}{M_u} * 100 = 3\%$	$A = \frac{S}{M_u} * 100 = 3\%$	$A = \frac{S}{M_u} * 100 = 83\%$	$A = \frac{S}{M_u} * 100 = 6\%$	$A = \frac{S}{M_u} * 100 = 3\%$
$A = \frac{A}{M_u} * 100 = 3\%$	$A = \frac{A}{M_u} * 100 = 83\%$	$A = \frac{A}{M_u} * 100 = 6\%$	$A = \frac{A}{M_u} * 100 = 3\%$	$A = \frac{A}{M_u} * 100 = 3\%$
$A = \frac{U_s}{M_u} * 100 = 3\%$	$A = \frac{U_s}{M_u} * 100 = 3\%$	$A = \frac{U_s}{M_u} * 100 = 6\%$	$A = \frac{U_s}{M_u} * 100 = 83\%$	$A = \frac{U_s}{M_u} * 100 = 3\%$
$A = \frac{V_u s}{M_u} * 100 = 6\%$	$A = \frac{V_u s}{M_u} * 100 = 3\%$	$A = \frac{V_u s}{M_u} * 100 = 3\%$	$A = \frac{V_u s}{M_u} * 100 = 3\%$	$A = \frac{V_u s}{M_u} * 100 = 83\%$

Table 4.35 Coercion acceptance Metric

This data identifies a sample group of 30 users who provided their views on the coercion techniques mentioned. TKT achieved a 83% rating which means it is level 1 , whereas ECG only achieves a 20% therefore giving it a Rank/Level 5 rating. This is done by finding out which areas the technique has the greatest value and then finding the corresponding level

4.5.7 Circumvention

Circumvention details the methods available to avoid the security measures in place or otherwise render them moot. Within coercion detection this corresponds to identifying if the sample *is* being coerced when it is not being detected or if the sample *is not* be coerced when *it is* being. There are a number of factors that make this unique compared to other aspects of biometric security. Coercion detection techniques consist of a number of different areas there are medical based data such as skin conductivity peaks; there are environment techniques such as proximity sensors; and there are user centric techniques such as tangible key techniques and whilst the first conforms to the liveness measures the others do not.

Environmental circumvention can range from: covering a data collection device with material designed to block the feed; removing the device from the relevant network; hacking

the system to prevent the data being provided to the decision maker etc. These threats are general to biometric security [156] and must always be considered. User centric circumvention is different as they are again based within both general security and specific biometric standards. For example including a fingerprint authenticator on a tangible key therefore only allowing the authenticated user access to the device. Therefore the following factors will be considered:

Spoof gathering

How difficult is it to gather a sample that allows a spoof to be developed, the lower this level the harder it is to gather the spoof. Therefore the following classification will identify this value (SGT = Sample Gathering Techniques):

1. Level 1 – 0 SGT
2. Level 2 - $\leq$ 1 SGT
3. Level 3 - $\geq$ 2 $\leq$ 3 SGT
4. Level 4 - $\geq$ 4 $\leq$ 5 SGT
5. Level 5 - $\leq$ 6 SGT

This will be given the letter 'S_g'in the following equations.

Accuracy

The accuracy of the technique is key, as it allows the system to correctly identify a user, therefore it is a key area for an attack to occur. The higher the accuracy the harder the system is to breach.

1. Level 1 – ERR $\leq$ 1%
2. Level 2 - ERR $\geq$ 2% $\leq$ 3%
3. Level 3 - ERR $\geq$ 3% $\leq$ 4%
4. Level 4 - ERR $\geq$ 4% $\leq$ 5%
5. Level 5 - ERR $\geq$ 6%

This will be given the letter 'A'in the following equations.

Linked Sample

As with liveness detection there is no necessity for the sample to provide any identification of the user instead the coercion samples are often separate from the authentication process. This can potentially cause techniques to become susceptible to the spoofing of data. Therefore linking the data more directly to the overall technique has the potential to make it harder for attackers to circumvent the system. Therefore this metric must identify the relevant linking between the different techniques, linked techniques will provide a more robust environment than non-linked.

1. Initial sample
2. Additional sample

This will be given the letter 'L_s 'in the following equations.

Sample Theft

Whilst permanence primarily deals with variations from within the sample which occur naturally, there is also the potential for samples to be stolen. This can especially be true for coercion due to the range of techniques which rely on easily stolen mediums such as keys, pass-codes or phrases. Therefore how easy it is to steal the sample must be highlighted and if there is a high risk of theft then this must be taken into account when considering suitable techniques for system integration.

Despite this importance, theft of coercion samples can be irrelevant, similar to liveness samples due to the lack of emphasis placed on authentication or verification. This is because most coercion techniques use data that is not in user specific, i.e. a pulse between 40 and 60 bpm, which can be the same for any number of users. Therefore if a coercion sample is stolen it does not have as great impact as the re-provision of this data sample would be sufficient for the user. Instead of going to the trouble of stealing data, the attacker could use their own sample data instead which would be much more easily created. For example a FACS technique does not require the same user as the authentication process as long as the correct AUs are being provided then the technique will work as intended.

This does not mean sample theft is not a problem within coercion detection, as this technique can have an adverse effect on certain techniques such as tangible keys. The theft of the sample is not the issue, it is the use of the key without authentication. The attacker is stealing the ability to provide a sample, and not the sample itself. With tangible keys this is a very important concept, as such keys would be easily available and cheap to manufacture and they would contain the most basic of technology. Therefore this metric will identify the

$$C = (S_g + A + L_s) * S_t$$

Table 4.36 Coercion circumvention metric

potential scope for spoof detection, mainly focusing on the effect this spoof data would have and not the applicability of spoofing overall.

1. Level 1 - No effect
2. Level 2 - Technique modifying effect
3. Level 3 - Technique breaking effect

This will be given the letter 'S_t'in the following equations.

Metric

This metric culminates in Table 4.36 which identifies the overall level of circumvention a coercion detection techniques has.

The main factor that effects coercion circumvention is the theft of a sample, therefore it is this value that is highlighted as the most impactful. Within this metric the sample theft level is used as a multiplier with the different levels denoting the effect of loosing the sample. If there is minimal or no effect then there will be no impact on the overall metric (X by 1) whereas if it will have an impact then the measurement will become increasingly poor.

To test this equation the data in Table 4.37 will be considered:

ID	Technique	S_G	A	L_S	S_T
1	FACS	1	2	2	1
2	Skin conductivity peaks	3	2	1	3
3	IFA	1	5	2	1
4	TKT	2	1	1	3

Table 4.37 Coercion circumvention testing characteristics

Considering IFA techniques: it is susceptible to 0 sample gathering techniques (S_G = level 1) , it has an ERR of 6.9% (A = level 5), it is an additional process to gather the sample (L_S = level 2), and it would have no effect if the sample was stolen (S_T = level 1). This would equate to a overall rating of 8, which would mean it is the second most secure techniques within the data set, with FACS being almost twice as secure due to the better accuracy. The rest of the calculations can be seen in Table 4.38.

Process	Algorithm
1.	$U = (1+2+2)*1 = 5$
2.	$U = (3+2+1)*3 = 18$
3.	$U = (1+5+2)*1 = 8$
4.	$U = (2+1+1)*3 = 12$

Table 4.38 Coercion equation calculation

4.6 Conclusion

Throughout this chapter the focus has been on developing the novel taxonomy for liveness detection, making sure that it would be viable for integration across other security sub-systems. There are a number of equations and algorithms that when combined will provide an overall classification that will be able to be given a specific value. This value will be comparable to the other values around it therefore providing a wealth of data that can help system designers, developers, and researchers in choosing the correct biometric liveness detection characteristic.

The taxonomy created offers what is required, it can take the different characteristics provided by liveness and coercion detection techniques, and provide an overall security rating which can then be used by developers, administrators research, systems etc. to promote dynamic security and development. One highlight of this research is the development of the coercion detecting techniques, this area is an extremely novel and original area as little research has been conducted in it.

Whilst the taxonomy appears to answer the objectives, to fully understand how it will work testing must be undertaken to discover the salient factors, the subsequent chapter will focus on these.

Chapter 5

Testing and Evaluation

> When new students tried an experiment that was particularly successful in terms of explosive force, the result was often a cross between a major factory refit and a game of Hunt-the-other-Kidney.
>
> Pratchett, 2004

A taxonomy has been developed to address the problems facing liveness standardisation. This taxonomy is to be applied to both liveness and coercion therefore the final stage is to test the taxonomy and analysis its effectiveness. This will be achieved by comparing current peer reviewed research data with the output from the taxonomy which will show if the expected output has been achieved. The final stage is to highlight the algorithm's effect on security and how the taxonomy uses the algorithm to create usable output.

5.1 Taxonomy Testing

The taxonomy testing is a key area of research as this section will highlight what practical application the taxonomy has. The taxonomy will provide the input data for the algorithm which will provide an overall level of security for the system. This value can then be compared to other systems and other variations to see what will occur when techniques are changed. To begin the testing process a selection of biometric devices will be considered. These are being used due to their popularity within biometric security, there acceptance and the ease of integration with liveness and coercion as identified by [144] and [103]. Whilst there are other techniques available these are the ones being focused on:

1. Fingerprint
2. Iris Scan

3. Facial Scan
4. Vocal Scan

   As well as these traditional approaches there will be a small selection of less applied techniques such as:

5. ECG
6. Gait

Each approach can contain different sub-styles such as the Dermalog [48] technique for fingerprints which contain the overall system as well as just the algorithm. Sometimes this will be tested multiple times as it can be applied to different areas such as the Dermalog system and techniques. This is because there are different types of liveness detection, however some are experimental and as such have very little data available e.g. [4], therefore whilst choosing only two aspect within each area will allow data rich examples to be used therefore making it easier to check the validity of the taxonomy.

The next section will discuss the taxonomy testing and will consider factors such as: does it work? Where are the problems? What can be done to improve it overall.

5.1.1 Liveness Application

Fingerprint Technique

For fingerprint devices the Dermalog algorithm will be considered as it can provide excellent speed, accuracy and reliability. There are two techniques in use: one of which is the Dermalog system and the other being just the standalone algorithm both of which were included in the LivDet 2013 Fingerprint Liveness Competition [48]. This competition attempts to find the best new liveness detection techniques for fingerprints in two categories: algorithms and systems. Within this competition the algorithm achieved an 84.63% accuracy rate taking into account FRR and FAR against a variety of tests including live, spoof and cadaver samples, along with a EER of 1.2%. The taxonomy outputs the overall level as 1.7: which equals the highest range within this taxonomy.

This information shows that the taxonomy outputs the same as the research completed by [62] [48], therefore proving that the taxonomy is working correctly. One thing to consider is whilst the system technique could be expected to be of the same taxonomy level, as the standalone algorithm, this is not necessarily the case. This is due to the additions a system requires, and the inherent complexity therein. For example universality is rated as 2.67 (V2) compared to the algorithms value of 2.33 (V1) this equates to a 13.6% difference.

Even more telling is the difference between the two techniques collectabilty values which equates to a 100% increase for the system based method due to the additional hardware requirements. This shows that the addition of hardware can have a dramatic impact on the overall effectiveness of a techniques.

Universality	Collectabilty
$\frac{V1-V2}{(V1+V2)/2} * 100 =$	$\frac{V1-V2}{(V1+V2)/2} * 100 =$
$\frac{2.33-2.67}{(2.33+2.67)/2} * 100 =$	$\frac{1.5-4.5}{(1.5+4.5)/2} * 100 =$
$\frac{0.34}{(5)/2} * 100 =$	$\frac{-3}{(6/2)} * 100 =$
$\frac{0.34}{2.5} * 100 =$	$\frac{3}{3} * 100 =$
$= 0.136 * 100 =$	$= 1 * 100 =$
13.6% difference	= 100% difference

Table 5.1 Dermalog classification differences

The main problem with the data collection for these techniques is that they are based on commercial techniques therefore some of the data is restricted, and as this system is also used in the American law enforcement AFIS installation, there is even less information on the technical features. However whilst this does cause issues due to the lack of transparent data there is still enough to run through the taxonomy and whilst the inclusion of the technique in AFIS restricts the data, it also validities that technique as a powerful liveness approach which should score very highly within the taxonomy. Therefore both the information gathered from [62] on the techniques accuracy and the validity and data gathered from [48] ratify the taxonomies results. The algorithm scoring the highest overall level and the system was level two because of the reliance on additional hardware [62] [183].

Facial Techniques

The next set of data to be considered utilises the work conducted by [112]: specifically the work on facial recognition based spoof defences. The universality metric is of an average level as the technique requires some minor hardware and software additions such as additional video cameras. Correspondingly this technique is robust to permanence altering standards as represented by the low permanence metric. This corresponds with the research, as the technique has been tested against 40 to 140 images with an overall mean FAR of 12.5% and a FRR of 3.06%, as shown in Table 5.2. This represents a lower level of security and accuracy

due to the high false accept rate within the system which corresponds to a performance of 5.47 and the circumvention of 3.12 both of which are amongst the worst tested within the taxonomy. Overall this technique is one of the worst scoring because of the high FAR and the poor universality of the technique as it focuses on a single demographic samples. If the technique was to look into a wider range of demographics then it could potentially improve the level of security dramatically.

Number of Users	FAR	FRR	Total
40	12.5	2.5	15
80	12.5	3.75	16.25
120	11.66	3.3	14.96
160	12.5	3.12	15.62
200	13	3	16
240	12.5	2.9	15.4
Median Value	12.5	3.06	15.51

Table 5.2 FAR/FRR for Facial Modality Data [111]

The next standard used to test the validity of the taxonomy was a fusion based feature and texture analysis conducted by [91]. This has been used for two reasons: firstly it utilises two important standards within facial liveness detection techniques which are static feature extraction (SFE) [39] and local binary patterns(LBP) [119][60]. Whilst these factors can be used individually, it is expected that when used together a more secure environment will be created as it can address more areas of concern.

This section will look at these three techniques, fusion, SFE and LBP, to identify if this hypothesis is correct or if the fusion does not achieve a higher level than the individual techniques. This highlights one of the key purposes of the taxonomy, to identify the best fusion factors to consider as well as identifying if fusion is relevant at all. This can then be used to choose the most appropriate technique for integration.

The initial data suggests that the above hypothesis is correct, and the feature and texture analysis fusion technique is substantially more secure than it's constitute techniques which produces an overall rank of 1.6 compared to the 2.1 for both the LBP and SFE techniques. This is a 27% increase in security as identified by Table 5.3 and is due to factors which would normally be difficult to identify. However with the taxonomy it is comparatively easy to identify what issues contributed to the level of security.

Security Difference
$\frac{V1-V2}{(V1+V2)/2} * 100 =$
$\frac{1.6-2.1}{(1.6+2)/2} * 100 =$
$\frac{-0.5}{(3.7)/2} * 100 =$
$\frac{0.5}{1.85} * 100 =$
$0.27027 * 100 =$
27.027% difference

Table 5.3 Difference between FBP, SFE and FTA

Table 5.4 shows that, according to the data, the fusion technique is superior to the separate techniques. This is seen as most of the classifiers for the FTA technique are superior to the two individual categories except for universality and collectability.

Technique	Universality	Permanence	Collectabilty	Performance	Acceptability	Circumvention	Total
Fusion[91]	2.67	0.67	2.00	1.47	1.00	1.53	1.6
Static Feature Extraction	2.67	1.33	1.67	2.00	1.00	2.08	1.8
Local Binary Pattern	2.67	1.33	1.67	1.80	1.00	2.03	1.8

Table 5.4 FBP, SFE and feature and texture analysis data

The universality of the techniques are all the same because they all share the same requirements and sample gathering techniques. This means that the individual standards can be used across the same amount of techniques as the fusion standard. Collectabilty has the same issues as it is reliant on the same form of biometric, if there was a large deviation with universality it would denote a change of biometric style.

The fusion aspects of the technique highlight the improvements that can be seen from the data, for example the improvement to permanence can be directly correlated to the application of the two techniques and the thorough testing and robustness of the algorithms discussed within [91] research. This research specifically incorporates facial occlusion detection and negation techniques, therefore improving the robustness against benign and medical noise.

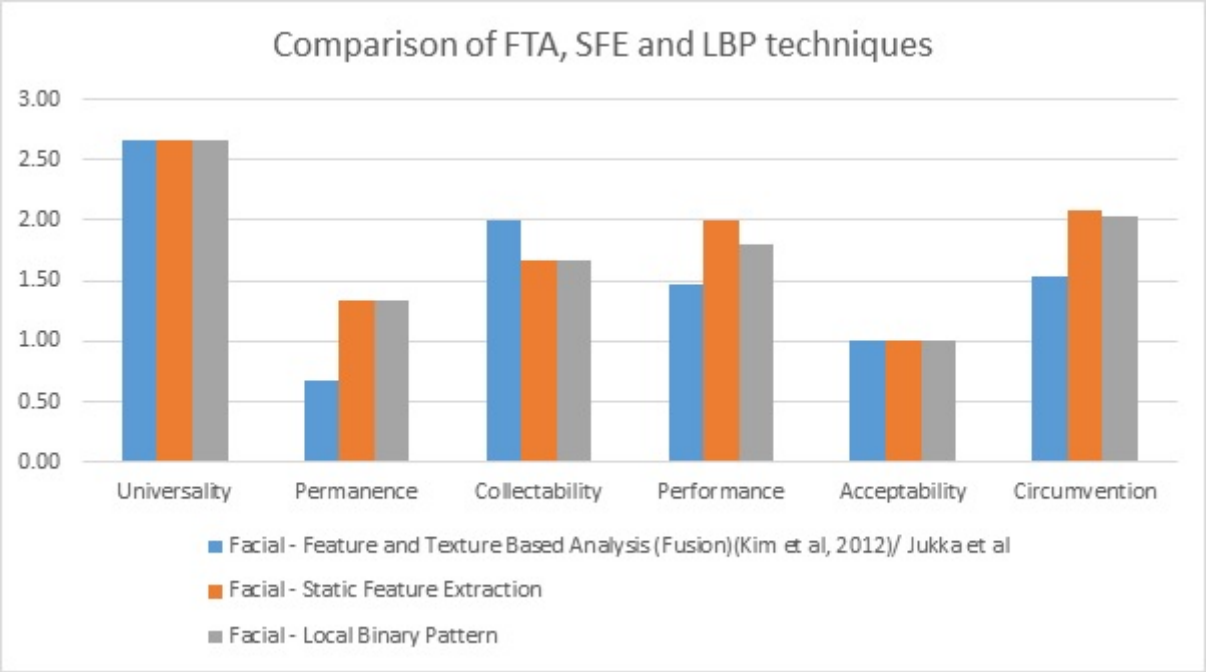


Fig. 5.1 Comparison of FTA, SFE and LBP techniques

As can be seen in Figure 5.1, collectability is the only factor to exceed the individual techniques and this is due to thee different techniques FTA can employ. As [91] identifies: there can be voluntary and involuntary techniques, therefore denoting samples that are not automatically taken during authentication, instead they are innate samples that are taken separately during the sample collection period. Due to this need for additional hardware the overall level of collectability is increased and the techniques suffers accordingly. The performance of the fusion technique is superior to both of the individual techniques which is due to two factors. Firstly the improved permanence of the fusion technique reduces the overall degree of intra-user variance by included the facial occlusion techniques and secondly the overall ability to spoof the technique is similarly reduced due to this increased difficulty of spoof data attacks. This corresponds to a 30.5% in performance compared to SFE and a 20.18% when compared to LBP as shown in Table 5.5.

One factor that has become increasingly problematic within this research is acceptability as it provides a very subjective view of a technique and is very difficult to measure. However it does produce relevant data for the system. The final area to consider is circumvention, and once again the fusion technique is superior to the individual techniques. As with other sections the primary factors that improves FTA, compared to SFE and LBP, is the permanence and performance of the technique as it is quicker to gather data, and more accurate primarily being due to the integration of facial occlusion negation characteristics [91] [52].

Difference between SFE and FTA	Difference between LBP and FTA
$\frac{V1-V2}{(V1+V2)/2} * 100 =$	$\frac{V1-V2}{(V1+V2)/2} * 100 =$
$\frac{1.47-2}{(1.47+2)/2} * 100 =$	$\frac{1.47-1.8}{(1.47+1.8)/2} * 100 =$
$\frac{-0.53}{(3.47)/2} * 100 =$	$\frac{-0.33}{(3.27/2)} * 100 =$
$\frac{0.53}{1.735} * 100 =$	$\frac{0.33}{1.635} * 100 =$
$= 0.305476 * 100 =$	$= 0.201868 * 100 =$
30.5476% difference	= 20.1835% difference

Table 5.5 Difference between FTA and SFE/LBP

From the analysis of both the taxonomy results, and the relevant research, it is obvious that the fusion based technique is superior due to the improved performance and circumvention. This result become apparent using the taxonomy to identify levels and was proved by the corroborative evidence provided within the research of both [52] [92].

Iris recognition

According to [177] iris recognition techniques are on of the most secure gathering standards within biometric authentication especially when compared with other, less robust, techniques such as vocal recognition which suffers tremendously with noise variations [114]. Therefore the expectation was that iris recognition techniques, when tested, would indicate that the technique was very secure, but some susceptibility to noise variations factors, especially medical based noise e.g. [137]'s cataract research as well as occlusion effects such as glasses, false eyes, contact lens etc. [177].

Figure 5.2 shows the sample data used and was taken from two specific pieces of research. The first considered the effects of contact lens occlusion which is a factor that can be very impactful as, according to the Association of Contact Lens Manufacturers, three million people in the UK and over thirty-one million people within the USA use contact lens [6]. This means that the impact contact lens can have when identification and gathering data can potential be large. This can cause problems for noise occlusion as well as ease of circumvention due to the familiarity many people have with the medium.

The second dataset used deals with the Purkinje images [96]. These images use data gathered from one of the four surfaces within the eye that can reflect bright light: the front and back of the cornea as well as the front and back of the lens. The z distance is then

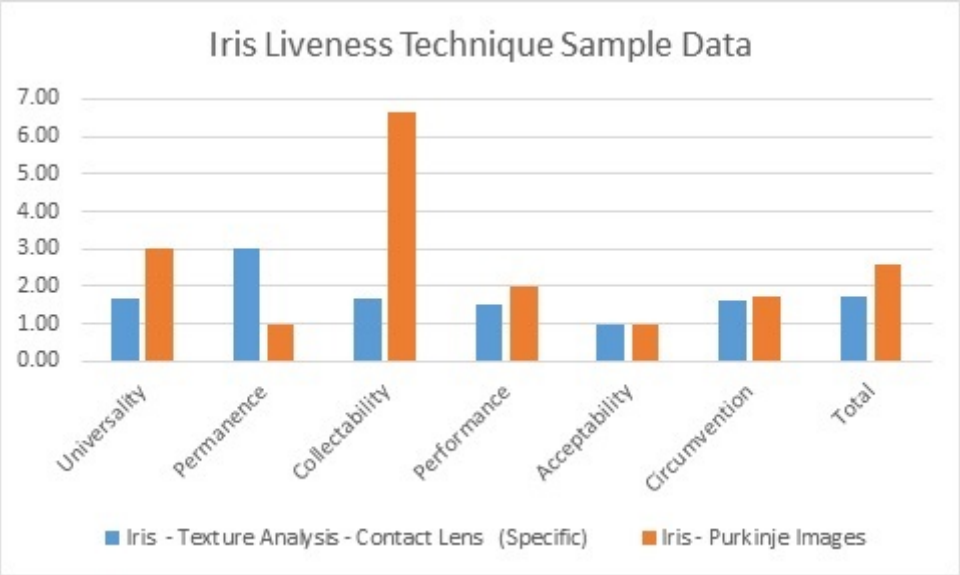


Fig. 5.2 Iris sample data

calculated to identify the liveness of a sample, alternatively other natural factors could be used such as cornea diameter and refraction rate as they would differ between a live sample whereas they do not in a false sample. These techniques are used within iris recognition and are covered in the work of [96] [89] [97] therefore representing a valid and accepted techniques of liveness detection.These two techniques represent the two main areas of iris based liveness detection: general detection and occlusion prevention. Whilst both consist of some similar concepts they have very different characteristics, something that is seen within the taxonomy results.

As Figure 5.2 shows there are some similarities however whilst the two techniques consist of the same biometric they are quite different. The main similarity focuses on the overall circumvention susceptibility of the techniques and is due to their overall similarity. However there are some differences to note, mainly due to Purkinje image's reliance on custom cameras and installations. Dude to this addition technology there overall implementation is less secure.

The rest of the classifiers are quite different and they equate to a substantial difference in overall security as shown by Table 5.6, once again the primary factor that causes this difference is the inclusion of additional hardware, as this effects multiple areas of the taxonomy.

The degree of permanence within both techniques is very similar however the benign susceptibility of the Purkinje image technique is less robust due to its reliance on very specific distances within the sample capture process [177]. The collectability of the technique again

Universality	Collectability
$\frac{V1-V2}{(V1+V2)/2} * 100 =$	$\frac{V1-V2}{(V1+V2)/2} * 100 =$
$\frac{1.8-2.6}{(1.8+2.6)/2} * 100 =$	$\frac{1.67-6.67}{(1.67+6.67)/2} * 100 =$
$\frac{-0.8}{(4.4)/2} * 100 =$	$\frac{-5/(8.34)/2)}{*} 100 =$
$\frac{0.8}{2.2} * 100 =$	$\frac{5}{4.17} * 100 =$
$= 0.363636 * 100 =$	$= 1.199041 * 100 =$
36.3636% difference	= 119.9041% difference

Table 5.6 Iris Technique Differences

is the same with one major deviation, the additional custom hardware needed to allow the technique to work as intended. This causes a substantial variance within the overall metric and causes a difference of 119.9% making it one of the most damaging factors when considering this technique as seen in Table 5.6. Therefore whilst the research indicates that the technique is viable from an accuracy standpoint the issues surrounding the collection of sample reduce the overall effectiveness. The additional hardware factor is the main damaging feature for the Purkinje image techniques as once again within both the performance and circumvention measurements additional hardware is needed. Once again the appropriateness of acceptability is questioned due to the lack of research in public acceptance of liveness sample gathering. Whilst it is very important overall for the success and security of a system: the lack of appropriate data on acceptance can cause data deviations to occur. Therefore it must be considered whether it would be best to include "assumed data" or ignore this metric in favour of returning with primary data at a later stage. Consequentially the overall results indicate that, except for certain areas such as susceptibility to noise deviations, the Purkinje Image technique is superior to texture analysis technique. The main flaw for Purkinje image techniques is that it is very specific in scope and therefore this must be considered when choosing the most suitable overall technique.

The taxonomy backs up the research findings once again showing the usefulness of the taxonomy. However one factor is becoming increasing obvious, the lack of a uniform measuring system within liveness data collection and testing is a major problem, as not all researchers discuss equally on different measurements such as sample acquisition time, FAR, etc. and even utilise different terms for standards such as FTE/FRR/FAR and EER. Therefore in addition to the advantages mentioned in earlier chapters, this taxonomy will also enable a

uniform measuring system within liveness sample collection and testing which will enable data comparisons to be more easily achieved.

Vocal Recognition

Vocal recognition is rarely used, except within fusion centric systems, due its innate poor security caused by the high degree of noise susceptibility. Therefore the main focus of this techniques is to develop fusion centric liveness techniques that can minimise this flaw. For these reasons a fusion technique was used to test the taxonomies robustness for vocal recognition, specifically the work conducted by [38]. This technique utilises lip motion which is intrinsically linked to audio and visual keys. This technique focuses on the fusion between two different biometric and liveness techniques whereas other liveness techniques such as the feature and texture based analysis described by [91] focuses on fusion of the liveness techniques only. This deviation is an important consideration as the liveness fusion will normally produce techniques that are overall more secure, whereas the biometric fusion will produce more universal techniques due to the innate reliance on different sample types.

As Figure 5.3 shows the overall rank for this technique is 2.81 and whilst this is not, comparatively, a high rank it does identify some interesting characteristics. This is a fusion based technique so there are a lot of additional factors to consider. The technique has comparatively few additional technology needs and can be used in a variety of different techniques which is indicated by the comparatively low universality rank in fact this is where vocal based techniques excel due to the innate universality allowing fusion to occur more easily than other techniques [46] [81].

This universality can provide a number of advantages mainly being the comparative ease of installation. This universality does have negative features which are primary highlighted within areas such as permanence and collectability. This is a trend that can be seen across a number of techniques, the lower the universality, therefore meaning the more universal the technique, the higher the permanence which can be seen in the work of [111] [177] and [38]. This is highlighted within Figure 5.4, the greater the universality the lower the permanence and visa versa. In practical terms this indicates that for a technique to be universal then minimising the susceptibility to benign and medical based deviations is paramount.

The collectability of vocal techniques is very similar to other techniques and has no obvious distinguishing features. This is ratified within the original research [177] [38] which once again shows that the taxonomy is working. Similarly the overall performance of the system matches a lot of the other techniques, primarily due to the speed with which the system can operate, and the lack of specific hardware needs, therefore keeping the technique viable due to the accuracy and speed. One further factor to consider is that this technique

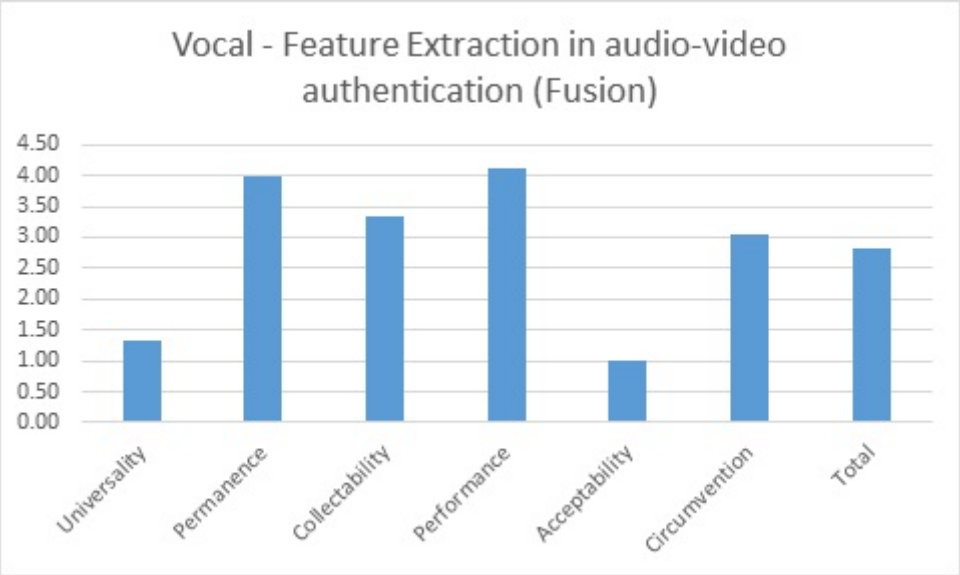


Fig. 5.3 Vocal technique analysis

can be easily combined with a profiling environment which would enable a much greater degree of integration and can help minimise some of the downsides of this technique, e.g. such as providing dynamic noise solutions when the environment calls for them. However this degree of integration, as mentioned within [38], could potentially cause other factors to be considered such as the addition of technology.

Novel Techniques

One of the defining features of biometric security is that there are new techniques being discovered and adapted all the time. Therefore the taxonomy will have to have enough flexibility to deal with this novel techniques. Therefore this section will consider some novel techniques of liveness detection will be considered alongside the more established techniques discussed earlier. A portable ECG technique will be considered as the technique itself is novel. The addition of mobility will allow the taxonomy to identify the salient attribute and to identify if the mobility factor detracts from the security factors. The focus of this testing is the work conducted by [152]. Normally a liveness detection technique takes a factor from the original authentication sample, or from a linked sample, and uses this sample to denote liveness. This separation is often needed as the biometric sample does not innately show liveness, e.g. fingerprints do not denote the users state they merely identify the user. However, this is not the case with ECG based biometrics as ECGs use the electrical representation of the heart as a way to uniquely identify the user and therefore automatically proves the users liveness as there is a need for a living heart during the sample gathering process. The main

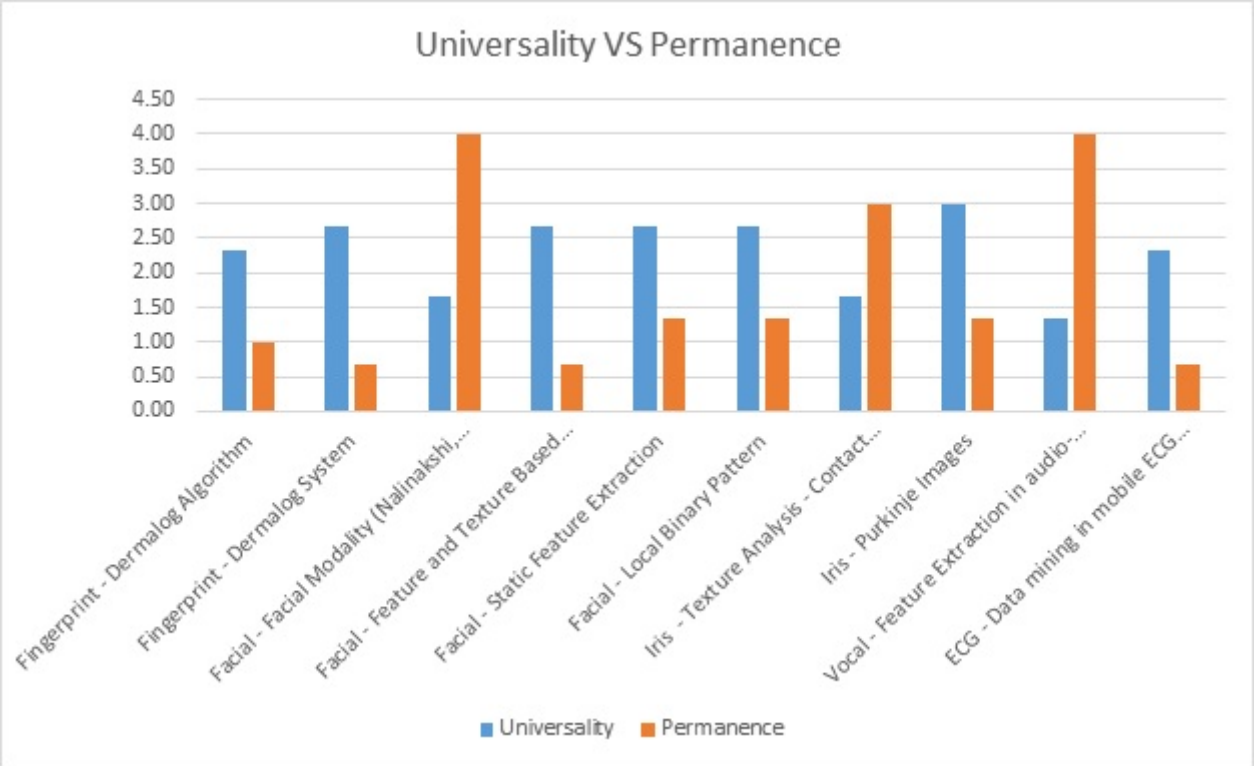


Fig. 5.4 Universality VS Permanence

problem with ECG biometric techniques has always been the reliance on additional hardware and software however this problem has been reduced by the inclusion of mobile ECGs. Even so the degree of hardware and software required, to gather the sample, is still very high when compared to other techniques such as iris texture analysis which only requires a basic camera. Therefore the main factors that the taxonomy will highlight will be the collectabilty issues alongside its very good accuracy and permanence characteristics.

Figure 5.5 shows the overall characteristics of this techniques and the first conclusion that can be made is that there is still a major problem within collectability. As, whilst it does not rely on the traditional ECG sample gathering system but uses a mobile system, it is still reliant on specific hardware and software subsequently equating to a high level of collectability. A solution to this would be to reduce the hardware reliance by creating a more universal ECG sample gathering techniques, one that is software based only, for example. This is identified within a variety of research including [114] [153] as well as being backed up by the taxonomy.

There are also a number of other conclusions that can be gathered by considering the results and comparing them to other techniques. Firstly the universality of this technique is very high as the technique can be used across all other liveness techniques, however the main

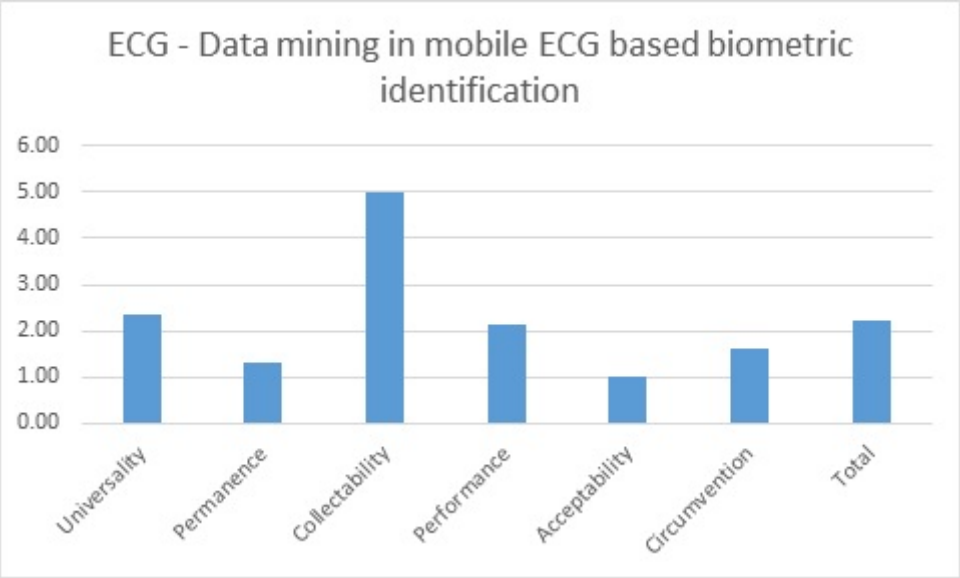


Fig. 5.5 ECG Data Mining in mobile ECG based biometric identification

problem will always be the integration of hardware and software. One major advantage of ECG that there are few benign noise factors that can affect the sample, therefore making its permanence very high. There are some factors that can effect sample collections such as powerful electronic presences, therefore a suitable location for ECG data collection, taking into account knowledge of geographical features such as the location of high power cables, would be sought. This is an important factor to consider especially for mobile ECGs as the emphasis is on the mobility of the device and whilst it is ideal that the operator knows where all interference nodes are located the addition of robust noise filters must also be used. Many of the same conclusions can be gathered surrounding the medical aspects of the system as well, as there are some minor noise variations that can occur due to the natural physiology of the body such as lung conductivity and heart rate variations [152].

The overall performance of this technique is positive because it is very accurate and very hard circumvent. This is partially due to the lack of research surrounding this technique and also due to the lack of spoof potential. As the technique is taking data directly from a user's heart, there is a minimal opening for spoofing data. Once again the prime problem focuses on the overall degree of hardware dependence, a trend that is becoming more and more prevalent throughout most of the techniques tested by the taxonomy.

5.1.2 Coercion Application

This section will cover the taxonomy testing focusing on coercion detection techniques. The coercion sample testing is different to liveness as there is a lack of primary data to test with

the taxonomy due to the novel techniques identified. Therefore whilst there are some current coercion detection techniques the data associated with them is sporadic they will still be used to identify if the taxonomy is functional for coercion detection. As well as the available data dummy data will be used that is associated with the novel approaches developed within this research.

Within this section the testing is split up into two categories firstly current techniques that have been discussed within research: secondly the novel techniques presented within coercion detection development section.

Tangible Key Technique

The first technique to be considered is Tangible Key Technique (TKT) which revolves around the use of a specific piece of hardware or software that is tangible (in its base form, or requiring additional hardware such as a phone in the case of an application). The main advantage of this technique is that the device is heterogeneous and therefore it does not discriminate against users, therefore it can be used within almost any technique. For example a fingerprint sample cannot function if the user does not have fingers which would cause collectability problems as well as discriminating against the user, however a TKT does not have these problems as the device can be easily developed and replaced. However there are some issues surrounding these devices: for example the tangibility can be stolen, lost or damaged, an app can be corrupted or the medium it is installed on can be stolen, damaged etc. [156].

As Figure 5.6 shows: the universality level of TKTs is comparative to other coercion techniques which is due to the requirements surrounding the additional hardware. This factor transcends coercion and liveness detection as one of the main factors to consider within the taxonomy as it has had the the most profound influence on each technique included so far. The major positive factors highlighted indicate that it is one of the most heterogeneous techniques within coercion detection due to its complete separation from any technique and therefore can be used with any other technique. For example a FACS technique is only relevant when conducting tests using the face i.e.facial scans or iris scans.However this technique is heterogeneous and very resistant to permanence altering factors.

This leads to the second advantage: the extreme robustness to variance of either medical or benign factors. As the technique revolves around a boolean factor, the key is either pressed or it is not pressed, then the degree of things that can impact the technique is minimal. Environmental conditions, unless extreme, will not affect them, ambient characteristics will also have a very limited effect such as light or humidity levels. The same can be said

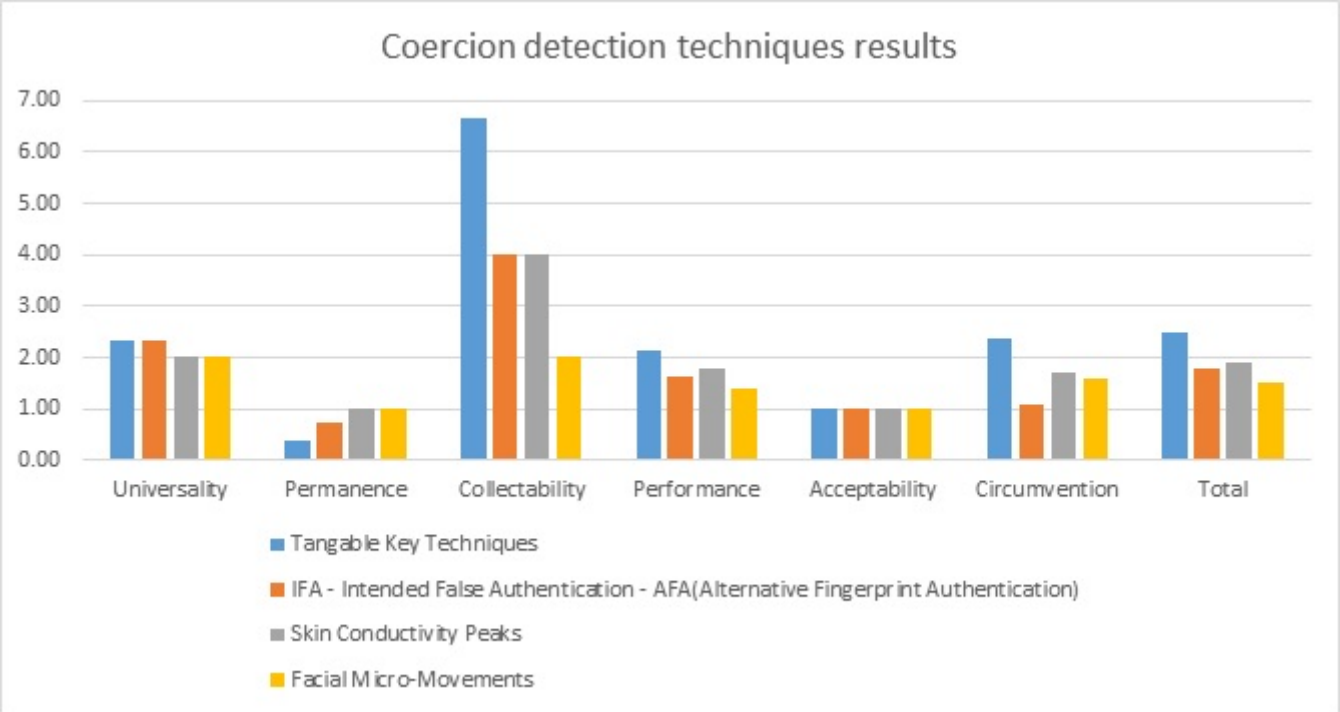


Fig. 5.6 Coercion detection techniques results

regarding medical variances as the user only has to press a button and only very serious medical implications will potentially cause any overall problems.

The collectability of this technique once again highlights the impact additional hardware can have within this taxonomy and once again shows how important it is. The performance of this techniques is the worst out of each coercion technique presented and this is due to the comparative ease of circumvention that can effect a TKT. For example the actual key itself is a very simple piece of hardware without any innate security therefore it would be comparatively easy for an attacker to create one themselves and then use this fake key. Factors could be put into place to improve this factor such as encryption, innate security, random ID etc. However with each additional security feature the overall technique will become less universal and will decrease the overall effectiveness of the technique, as well as questioning the appropriateness of the technique. Therefore the performance, whilst when used correctly, is very accurate and quick, it is prone to this circumvention flaw which drastically reduces the security level when compared to the most secure technique, FACS. Table 5.7 shows that there is a difference of 46% between the two techniques which means the TKT technique is almost half as secure as FACS.

The taxonomy shows, whilst this technique has some applicable features, there are a lot of problems and questions surrounding the security, performance and ease of collection of

Security level difference between TKT and FACS
$\frac{V1-V2}{(V1+V2)/2} * 100 =$
$\frac{2.4-1.5}{(2.4+1.5)/2} * 100 =$
$\frac{0.9}{(3.9/2} * 100 =$
$\frac{0.9}{1.95} * 100 =$
$0.461538 * 100$
46.1538% difference

Table 5.7 Difference between TKT and FACS

TKTs. Therefore whilst there overall score is average the individual factors tell a different story, therefore highlighting another positive factor of this taxonomy, by classifying the different sections it is possible to identify what areas are most important. Therefore provide a thorough identification of techniques, not merely an overall final output. This flexibility allows for a more robust knowledge to be developed therefore helping the researchers and developers of the future.

Skin Conductivity Response Tests

Skin conductivity response signs can change when subject to strong emotions. The physiological signs of these emotions provide the information that is needed to detect the state of a user, in the case of coercion detection, these emotions would focus on negative ones such as fear and disgust[68] [160]. As Figure 5.6 shows the overall technique is secure, rating a level 2, which is superior to the TKT techniques and of the same level as the other two techniques. However it is not quite is secure as the other two overall. Considering the universality of the techniques, the main factor is that this technique is very reliant on additional hardware however it does provide a high universality and good applicability to other techniques across both liveness and security samples, as long as the technique requires contact with the skin. The permanence of this technique again is comparatively similar to the other techniques however there are some minor variation factors to consider. What effect do cosmetics/skincare products have on the degree of conductivity, if any? What medical factors can change the test including medication ingestion and generic medical variations

[53]. Once again the addition of hardware is one of the main issues when considering the collectability and negatively effects security level.

The main advantage of this technique is that it is accuracy and whilst the other technique also have high accuracy, this techniques maintains this accuracy across security, liveness and coercion fusion levels. However this technique has one caveat as it is dependent on background research and whilst this physiological testing has been considered for a number of years the application within coercion detection is very new and therefore it would be necessary to take every caution when integrating the technique.

Security level difference between TKT and FACS
$\frac{V1-V2}{(V1+V2)/2} * 100 =$
$\frac{1.45-1.09}{(1.45+1.09)/2} * 100 =$
$\frac{0.36}{(2.54/2} * 100 =$
$\frac{0.36}{1.27} * 100 =$
$0.283465 * 100 =$
28.3465% difference

Table 5.8 Circumvention level difference between SCP and IFA

This technique provides a good security level and whilst it may not be as secure as some of the other techniques, for example as Table 5.8 shows there is a 28% difference between SCP and IFA due to the already present use of this technique within other disciplines, the comparative integration could well provide a more valuable option to a system developer.

Intentional False Authentication

The first of the two novel techniques is Intentional False Authentication (IFA) which checks if a user is being coerced by allowing them to provide a sample that is deliberately incorrect. For example instead of using the index finger to authenticate the user will use a designated different finger or thumb and when this is detected the system automatically knows that the users is being coerced. This is identified as one of the most secure techniques. This technique can be applied to almost all other biometric and liveness techniques, as the user has merely to designate a separate authentication sample to be registered as the coercion measure. This

can be done by the user or by the system and has a number of advantages for example there is a limited need for additional hardware as the sample being used is from the same type and the heterogeneity it affords is excellent as most samples will have an easily identified alternative that can be used as depicted within Table 5.9.

Technique	Alternative
Index finger	Little finger Multiple fingers
Left iris	Right iris Blinking
Left palm print	Right palm print
Vocal key phrase one	Vocal key phrase two

Table 5.9 Examples of IFA techniques

There are some innate considerations specific to this techniques as the onus of this technique is reliant on the user instead of the system. This injects an additional factor, how subtly can the user present the IFA data to the system, for example when using vocal phrases this can be quite difficult because of the overt way the technique works, alternatively when using fingerprint techniques it is simpler to obfuscate the false data. Regardless this technique is still very resistant to permanence altering factors, which remains one of its primary strengths. Another advantage is that this technique does not rely on additional hardware which means it is less prone to permanence and collectabilty altering factors. The performance of this technique is very high and it is accurate but does cause the system to consider twice the samples during matching.

Due to the flexibility of this technique, the circumvention and spoof susceptibility of this are not as damaging but are still important to consider. The development of spoof samples for this technique is irrelevant as a spoof samples would produce no effective features. The way to spoof this sample would be to prevent the user's ability to provide the specific IFA sample but as one of the main underpinning features of this technique involves the users subtly then this is issue can be negated. Otherwise the general circumvention classifier is very low showing an excellent robustness to attacks. This is due to the innate attributes of the sample as they are directly linked to their parent sample and use all of the same hardware and software therefore once again relying on the official technique.

The main factors that have been analysed from the taxonomy data, is that this technique has the potential to be very powerful, it is resistant to permanence altering factors, circumvention techniques and has a good performance. One factor that is not positive is that it seems

that the best integration of this techniques would be within a fusion based system as it would allow a dynamic environment to be developed. For example a fingerprint/iris recognition system using an IFA coerced detection technique would be able to provide users with either an IFA fingerprint or iris sample and this could be adapted depending on the overall system requirements and features.

Facial Micro-Movement

The final technique is Facial Micro-Movement which is based on FACS. This is a novel technique that is based around the detection of emotion and the corresponding physiological characteristics therein. Within coercion detection the obvious focus is on negative emotions and ones that can be associated with coercion detection, mainly fear, anger, distress etc. This technique is based on the FACS which has been used within the affect testing area for a number of years and has been identified as the superior technique in the area [54] [78]. Therefore its integration into coercion detection is a reasonable jump and by testing it with the taxonomy it should demonstrate its potential and highlight any problem areas. FMM has a very high degree of universality as there is minimal additional hardware needed and only some additional software to help decipher the AU (action units), however whilst the technique works well with facial based techniques it obviously has no connection to other styles, therefore demonstrating a low universality.

The permanency of this technique is good due to the lack of intra-user variation, however there are some effects that must be considered. Whilst there are few traditional noise factors that can affect a FMM system, one of the few would be light levels as bright or low light may cause the user to change their AU accordingly, such as squinting in bright light. The other non-medical factor that could impact this technique involves the way the AUs are interpreted. As this FMM uses the FACs the methods of coding can be interpreted in different ways. The collectabilty, unlike other techniques, is good due to the lack of additional hardware reliance, the comparative speed of authentication and the adequate degree of universality. This provides a vastly superior collectabilty level compared to other techniques equating to a 66% at best and 120% at worst difference as show within Table 5.10.

The performance of this technique is dominated by the low level of sample spoofing. This is because the correct allocation and interpretation of AUs is all that is relevant. Therefore it could be the user supplying the data or it could be an attacker therefore making moot any potential security factors. This would mean that anyone could supply the coercion samples and it would allow the user to work throughout the security process. One potential solution would be to create a more linear process which merges authentications into the coercion and liveness sample use, however this creates its own problems which must be considered.

Universality	Collectabilty
$\frac{V1-V2}{(V1+V2)/2} * 100 =$	$\frac{V1-V2}{(V1+V2)/2} * 100 =$
$\frac{1.67-3.33}{(1.67+3.33)/2} * 100 =$	$\frac{1.67-6.67}{(1.67+6.67)/2} * 100 =$
$\frac{1.66}{(5)/2} * 100 =$	$\frac{-5/(8.34)/2)}{*} 100 =$
$\frac{1.66.8}{2.5} * 100 =$	$\frac{5}{4.17} * 100 =$
$= 0.664 * 100 =$	$= 1.199041 * 100 =$
66.4% difference	= 119.9041% difference

Table 5.10 Coercion collectability difference

Overall this is shown as a powerful technique with a lot of potential, but with one major negative, something that must be addressed to get the most out of the system. However once again the taxonomy enables these factors to be conducted and identified showing the worth of the overall taxonomy.

5.2 Algorithm Development

Whilst the taxonomy can provide plentiful datasets for comparison and analysis this was not the final goal for this research. The taxonomy creates informative values that are components of the final algorithm, which can be used as a measure for overall system security. However this is not effected solely by the taxonomy values, in-fact there are a number of other factors that impact the applicability of liveness and coercion standards within a security installation. The goal here is to create an algorithm that can produce a single output that denotes total system viability for different security techniques. It will contain factors that will be specific to the installation environment and will not change, unless a different environment is chosen i.e. different location, company etc. The coercion and liveness standards will change the overall algorithm, therefore allowing different combinations of techniques to be tested and the multi-layer fusion will denote the statistical security level within the installation. This can then be used to evaluate and analyse other liveness and coercion techniques.

Therefore the following section will detail the algorithmic development, components and testing.

5.2.1 Justification

Whenever a security feature is added to a system the requirement is that it solves a particular problem, and its successes depends on success of this task. Therefore it has its own inbuilt testing metric, if the requirement has been satisfied then the feature is a success and whilst other systems may include evaluation systems of their own making, the main focus is the efficiency and effectiveness of the new technology. When considered biometric, liveness and coercion techniques the taxonomy provides a method of comparative critical analysis. However it does not take into account the fusion factors within each area, this is where the algorithm will appear. The algorithm will take the liveness and coercion techniques, alongside a variety of other metrics, and calculate the overall security rating an installation has. This can then be easily changed to identify different combinations of techniques and the difference between overall classifications can be highlighted and compared, due to the inculcation of interval based measurements.

One of the first problems encountered when developing this algorithm was the quantity of components. The initial factors shown in Table 5.11 takes the metrics developed within the taxonomy section. However it does not make a lot of sense, as certain factors required additional calculations, especially the D_r component.

$$\sum_{i=0}^{n} y_i - \bar{y}^2 = A_s = \frac{t}{P/A_b} * D_r = \frac{A_s}{L_t + L_c}$$

Table 5.11 Initial algorithm

The initial idea was to include a number of authentication times and use this as a value. However this was not suitable s it did not take into account the potential of multiple authentication attempts a single user could conduct therefore some form of differentiation regarding scale was needed. The algorithm was split up, with the different components containing sub calculations that provided the final algorithm with only the required data. Subsequently the first step is to identify what factors the algorithm are vital as these components will identify different stakeholder groups that can impact the security within a system, and are catered to biometric security specifically, culminating in Table 5.12, the final algorithm.

$$A_S = \frac{T}{\frac{P - A_b}{P} * D_r}$$

Table 5.12 Final algorithm

Within the algorithm the purpose of these metrics is to highlight different variables which are time, users, device redundancy, liveness and coercion detection. These will be discussed in detail in the following section. These factors will produce an output 'A_S'which will denote the overall level of security for the current system set-up.

5.2.2 Algorithm components

This algorithm will have a number of uses, the main one being a method for a researcher or design to simulate a security environment and identify what techniques would work best together for a specific environment. This fusion measurement tool would improve biometric security development as it would allow a more thoroughly tested system to be developed without the same degree of expense. A secondary factor: this algorithm could be used within a automated system, as the values entered into the algorithm could be automatically edited and the output could be used within a context aware environment to provide the most secure security options the system has access to.

Time

Time has been used because it is a good way to detect the accuracy of a system, and is necessary to provide much needed context to the other components. For example if the algorithm output is based on one day only, then there may be problems that arise when it is not that day, or when the measurement is a longer or shorter value. The time value is mutable and can differ depending on the level of security and the expected outcomes. Small scale time periods may be fine for initial tests between 5-10 hours etc. However if the time value is desired to be longer, greater than 24 Hours, then other factors must also be conspired such a noise, user-base size etc.

The purpose of time within the algorithm is to create a snapshot over a specific amount of time, the longer the time the more susceptible the system will have been to security threats and problems, therefore increasing the time will correspondingly increase the overall calculation. The goal of the algorithm being a set of techniques that minimise the final result of the algorithm whilst still using the required time value.

One variation to time would be attempts to access the system. This may be more suitable for a number of environments as users may log on a number of times a day, therefore getting average data may become more viable. Instead of checking over a period of time for example one day, the metric would be for every one thousand login attempts. If this metric does transform into attempts to access then an increase would denote a potential reduction in security, as a FRR of 0.01% is much more likely to fail if it is tried 10 times instead of just

once. The greater the attempts to access, whilst maintaining a low security score, identifies a very secure system.

Within the algorithm time is denoted by the value T which equals time taken during authentication, therefore the more authentication attempts that occur the higher the time value will become. This value takes the amount of authentication attempts and combines it with the time taken, therefore the greater amount of authentication will denote a higher value, however as these values are divided by the Tn component, with N being the amount of departments/samples within the calculation the value will always be a manageable number. Therefore an example of this is shown in Table 5.13.

Time Calculation
a_i = number of people within department. l_i = logins per department k = constant value less than or equal to one. t_n = amount of different summation samples each equalling n+1
$T = \frac{\sum_{i=0}^{n} \frac{a_i}{l_i+k}}{t_n} =$ $T = \frac{\sum_{i=0}^{n} + \binom{16}{5+1}\binom{25}{30+1}\binom{10}{15+1}\binom{40}{50+1}}{4}$ $T = \frac{9.2}{4} = 2.3$
$a_i = 16, 25, 10, 40$ $l_i = 5, 30, 15, 80$ $K = 1$

Table 5.13 Time calculation process

This value allows the inclusion of both time taken to authenticate and the amount of times authentication occurs. Obviously the higher either of these values the less secure the system will be. This is because when more people authenticate the risk for a threat to occur is increased, and the longer it takes to gather the sample the easier it is to spoof. After the time has been identified the next component to identify is the user-base

User-base

The user base is the area most prone to deviations. This component identifies the quantity of users being parsed throughout the system during a specific time. This metric can change dramatically depending on date and time, for example: when considering frequency and

number of logins per day as well as the user base, an initial assumption would be that the more users within a system then the more logins will occur. Whilst this is a reasonable assumption, other situations can impact this dramatically. For example it would be reasonable to expect 10 users to have less authentication attempts then 20 users. However if the 10 users are power users, and log on 30 times each within a particular session equalling 300 authentication attempts and the 20 users log on 10 times each equalling 200 authentication attempts, it shows that the values alone do not show all of the picture. This highlights the importance of user location and composition when considering what to expect from the user-base

Therefore the purpose of the user-base is to identify the style of access attempts, when will the access happen, and to identify what specific variations there are within the user base. A broad generalisation will be identified for the user base assuming that everyone can use the security based information – e.g. fingerprint scanner, with temperature based liveness detection.

To calculate this value the first thing that was identified is the amount of potential users within a installation. For the purpose of this example a university is used: there would be three categories of user, each representing a different organisational group within the university,a department, a school which comprises of one to three departments, and a faculty comprising of 'n' schools and departments. As there can be a variety of values within each area the mean of users will be found which will enable a more regulated calculation to be developed. The mean is being used as it is almost impossible to predict, without additional tools, how many authentication attempts there will be on a system at any one time period, as there are many reasons that may stop users from accessing a system, such as illness, holidays, meetings etc. Table 5.14 identifies a faculty level calculation which contains seven departments.

Table 5.14 shows the mean assuming that 'n' is the quantity of users within a department, and A_n equates to the sequence of users:

This would then be used to create the sum of the users within the simulation as shown by Table 5.16:

The final stage would be to calculate the final mean as shown by Table 5.17:

It may seem unnecessary to include the sigma summation for this simple equation and whilst in this example the data set is very limited, if implemented into a large enough organisation the amount of data may become extremely large and therefore become difficult to easily manipulate. Within this dataset the mean value for users is 25 this can then be used to show how many users are authenticating within the system, and is the first stage of creating

Users (an)	Attempts to Authenticate (aj)
16	20
25	100
10	10
40	50
55	200
10	40
18	10

Table 5.14 Mean data source

$$A_n = 16, 25, 10, 40, 55, 10, 18$$
$$\sum_{i=1}^{n} A_n$$

Table 5.15 User sequence

valuable data that can be used within the system. One potential use of this data would be to incorporate the results into a context aware environment.

This would allow the autonomic environment access to important data regarding user log on trends. To gather further data from this initial component additional factors would need to be considered such as the mean and absolute deviation. If Table 5.18 is used as the base level then Table 5.19 shows the absolute deviations and Table 5.20 the mean deviation. This can then be used to find out the mean divergence, therefore allowing the algorithm to check if a department is varying too much from the mean divergence, which could indicate a system error, lack of resources, or a potential security issue.

Whilst this data is useful from the algorithm standpoint, simply understanding the quantity of users is not sufficient, as it is unlikely that users will only login once a day, indeed depending what the environment is the amount of logins may contain huge divergence intra-departmentally. Therefore the next stage of development would be to identify the amount of logins per time period, this could then be combined with the amount of users to produce a valuable metric which will identify the amount of attempts to authenticate over each time period.

To do this: the average number of attempts will be gathered, this information can be used within either the time or the user component of the algorithm. Table 5.21 follows the same equation process as Table 5.16 and Table 5.17. Firstly users need to be identified, secondly

$$\sum_{i=1}^{n} A_n = a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7$$
$$\sum_{i=1}^{n} A_n = 16 + 25 + 10 + 40 + 55 + 10 + 18 = 174$$

Table 5.16 Sum of user simulations

$$\sum_{i=1}^{n} A_n = 16 + 25 + 10 + 40 + 55 + 10 + 18 = \frac{174}{7} = 25$$

Table 5.17 Mean of user simulations

average attempts to login/authenticate within a time period e.g. hours, days, months etc. Once again the data for this is held in Table 5.14.

Now that both of the means have been calculated the next stage is to generate the product of the means.

After these means have been gathered and the product of both is calculated the average authentication attempts for the particular time period will be highlighted, in this case one day in a university faculty. As shown in Table 5.22. The main problem here is that time is static and the user data is only based in a single area. To make this more useful for the full algorithm, more output is needed, potentially include multiple times/multiple departments. Currently all that can be gathered from the data is how many average attempts a time period there are. This then needs to be combined with other aspects for the algorithm to put these amounts into some context e.g. liveness detection FAR of 0.01 means that if this number is too great it will cause authentication problem.

This can then be used to calculate a threshold value which can denote if a system is, statistically, threatened. This is done by identifying the amount of authentication attempts and by finding out different divergences such as the mean divergence of average authentication attempts. This would then allow a systems administrator, autonomic element etc. to take advantage of this information, by providing more or less system resources nodes etc. This value is represented in Table 5.23:

The output of this calculation, R_a, shows how effective a technique inclusion will be. If a low score is identified then an administrator or autonomic element can be informed and any potential improvements can be deployed. This can also be used to represent a median deviation on access, if there is a continual low security identifier then this may represent a compromised environment. Finally this will enable the system to identify potential weak spots within a security environment. e.g. users normally log on with technique A, if users

Actual value	Distance from mean value (25)
16	9
25	0
10	15
40	15
55	25
10	15
18	7

Table 5.18 Mean variation

$$\sum_{i=1}^{n} x - \theta = 9+0+15+15+25+15+7 = 86$$

Table 5.19 Absolute deviation of user simulations

need to log on with technique B, due to impairment or technical issue, then the algorithm will endeavour this factor.

Therefore within the algorithm, after T has been calculated P must be found, and the following equation finds out the value of 'P' which is participants, which in turn denotes the user data for testing using the following value ranges:

To begin the initial value must be equal to or greater than 0, then the collective value of authentication is gathered by identifying the average amount of logins conducted using the same techniques as above with the constant value being included at the end of the equation.

The reason for this distinction is that whilst a user may log in only once a day within department X, department Y may face a huge amount of logins due to different classrooms, installations, subject etc. The summation calculates this value and makes sure that it is greater than zero.

Liveness detection and Coercion Detection

Like biometric samples both liveness and coercion techniques have different salient features therefore providing the overall algorithm with the information from the taxonomy allows different techniques to be identified and compared. The purpose of this metric is to provide a method of measuring the effect liveness and coercion brings to a system. When combined in the algorithm the higher the level the more secure it is, however it also will allow the inclusion of different options, e.g. needs high security but has very little space for additional

$$\frac{\sum_{i=1}^{n} x-\theta}{N} = \frac{9+0+15+15+25+15+7}{7} = 12.28 = 12$$

Table 5.20 Absolute deviation of user simulations

$$\sum_{i=0}^{n} \frac{A_n}{n} = 16+25+10+40+55+10+18 = \frac{174}{7} = 24.85$$

$$\sum_{i=0}^{n} \frac{A_n}{n} = 25$$

$$\sum_{i=0}^{n} \frac{A_j}{n} = 20+100+10+50+200+40+10 = \frac{430}{7} = 61.42$$

$$\sum_{i=0}^{n} \frac{A_j}{n} = 61$$

Table 5.21 Login attempts

hardware, therefore a software/intrinsic solution, whilst not as robust from a security view, may be a better solution.,

Biometric devices can be defined by a number of characteristics and a plethora of considerations must be taken into account when deciding on a suitable device for a system. There are a number of stages to consider beginning with what device is the most appropriate for the situation. The situation can have a number of requirements such as where the device is located, what is it being used for and for what degree of security will it be used i.e. is it a low or high security area. This will also consider what liveness and coercion techniques should be used, utilising the data gathered from the taxonomy as a starting area. This will allow multi-modal and multi-fusion factors to be considered.

Users	15	16	13	8	10	
Login attempts	15	25	20	17	8	
Liveness Technique	Coercion Technique					
1.47222	2.888889					
2.18055	1.548611					

Table 5.25 Algorithm company example

The purpose of this component is to identify the different biometric devices and techniques that work best together which can be based around multi-layer fusion e.g. liveness and coercion. This component will have a number of sub calculations that will provide the final device redundancy value for use in the main algorithm.

$$\sum_{i=0}^{n} \frac{A_n}{n} \cdot \sum_{i=0}^{n} \frac{a_j}{n} = 61/25 = 2.4$$

Table 5.22 Mean faculty data

$$\sum_{i=0}^{n} \frac{A_n}{n} \cdot \sum_{i=0}^{n} \frac{a_j}{n} = R_a$$

Table 5.23 Basic autonomic controlling algorithm

The device redundancy (D_r) is the metric that utilises this liveness and coercion data. This provides a level of security with a lower values denoting a secure system, whilst higher values denoting less secure as shown in Table 5.25. Practically it is possible for this value to be 0 which would mean that there are no liveness or coercion techniques included. If this occurs than the values being zero will have a negative overall effect throughout the algorithm. Therefore to prevent this problem occurring a constant is chosen is added to the sub calculation which is the standard exponent equalling 2.718 as shown in Table 5.26:

Device Redundancy equation
$\sum_{i=0}^{n} \left(\frac{l_d + c_d}{2} \right) + e^1$
$D_r \sum_{i=0}^{n} \left(\frac{2.5+0}{2} \right) + e^1 = 3.968$
$D_r \sum_{i=0}^{n} \left(\frac{2.5+0}{2} + \frac{1.33+3.33}{2} \right) + e^1 = 6.298$
$l_d = \{2.5, 1.33\}$
$c_d = \{0, 3.33\}$
$e = constant$

Table 5.26 *Dr* = Device Redundancy equation

This would prevent the lack of liveness and coercion standards causing failures within the system and would represent a single liveness and coercion technique being used, however there is the potential that fusion techniques could be encountered therefore the algorithm must be robust enough to support these factors. This is why the summation factor is used instead of a standard x + x. if this data was used then the identifying the peaks and troughs

Participants Calculation
A_i = number of people within department. L_i = logins per department K = constant value less than or equal to one.
$$P = 0 \leq \left(\sum_{i=0}^{n} \frac{a-i}{l_i} \right) + k \leq 1$$ $$P = 0 \leq \begin{pmatrix} 16 \\ 5 \end{pmatrix} + \begin{pmatrix} 25 \\ 30 \end{pmatrix} + \begin{pmatrix} 10 \\ 15 \end{pmatrix} + \begin{pmatrix} 40 \\ 80 \end{pmatrix} + k \leq 1$$
$P = 30.53$ $a_i = 16, 25, 10, 40$ $l_i = 5, 30, 15, 80$ K = time in hours (could be minutes, seconds etc.)

Table 5.24 Participants equation

would be important. Figure 5.7 shows the relationship between time elapsed and security rating:

There is a sequentially improvement in security performance as the time progresses however there is a large jump towards the end of the 12 hour segment, this is not due to any problems within the algorithm instead this indicates that there are issues with the data. Therefore to identify this anomaly the data has been analysed and it became obvious that the problem revolves around the user login and attempts data, specifically that there are fewer login attempts on certain days when compared to the amount of people who can log in. This is interesting as there are a number of potential ramifications here: firstly, and most benignly, is that the full compliment of users are not utilising the system. This could be due to a number of factors such as staff illness, holiday, and lack of need etc. and whilst this does impact the system overall it does not represent a threat innately. It does open a potential threat vector: if the normal login quantity is not being reached then the addition of a nefarious user will be harder to detect. The second factor could be due to a security breach within the system that is preventing legitimate users access, concepts such as poisoned cache attack, comprised password databases etc. would prevent system use becoming a DoS based problem. The third option is that there is a legitimate fault with the system, or there is an administration based attack.

As Figure 5.8 shows when the 'D_r' data changes, the overall graph changes as well. There are no sequential curves, however one factor that still stays the same is the anomaly towards the end of the time period. This can change dramatically depending on the liveness and coercion standards used therefore it must be identified early what factors are going to be

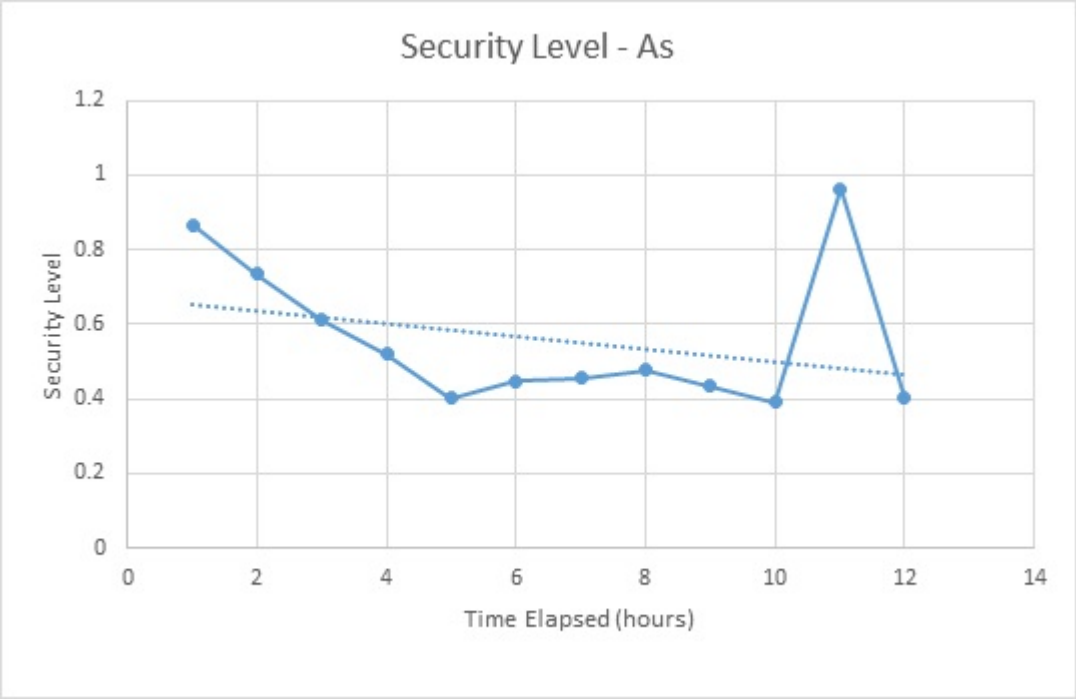


Fig. 5.7 Time elapsed vs Security

tested. For example if four departments are considered, the building they occupy is moving to a biometric security system but the developers cannot decide whether to include facial, iris, or fingerprint system. They understand the necessity of liveness and coercion technique and want to include a fusion techniques that makes use of two different methods. They decided that the measurements need to be in hours, as the company runs both at day and night and wants to see what the difference is between the two.

This data represents a high liveness-low coercion and low liveness - high coercion installation.

As Figure 5.9 and Figure 5.10 show the overall difference between different techniques can change the overall algorithm output by comparative small amounts. The cause of this small change is the lack of data. The more data used within the system the more accurate and responsive the algorithm will become, therefore highlighting the necessity to develop a database of samples that can be used for overall understanding of the different security increases. These graphs show that there is a sequential improvement as time progresses within the system, and this could be because of the stability improves over time, the lack of burst data transfer amongst others. The one problem once again is due to the different data identification, with less logins than users as highlighted above.

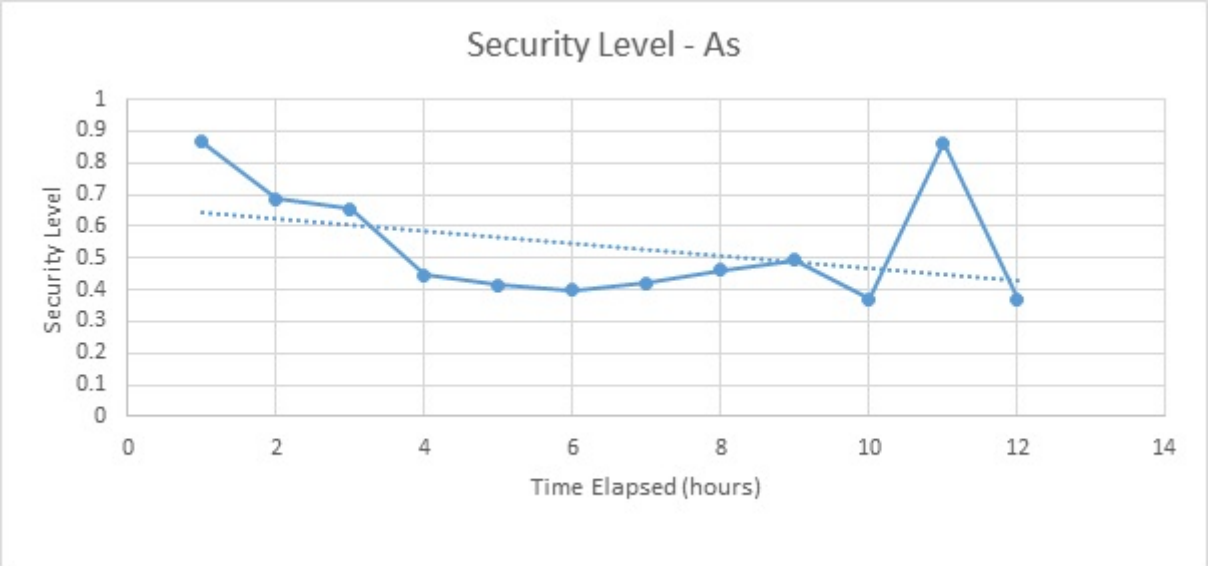


Fig. 5.8 Security level over time

Anomalous User-base

The permanence and collectability of biometric samples is a key factor within the taxonomy and whilst this has a deep impact on the individual techniques, it is also a factor that needs to be highlighted within the overall algorithm. If a user is unable to use a biometric due to an impairment then this must be taken into account when designing the systems security. Within the taxonomy the permanence and collectability measurement highlight what could affect the techniques. Whereas in a security installation actual data can be gathered and therefore can be incorporated into the overall design. For a designer or developer the data can easily be gathered by working with the companies human resources department, however as a sample base Table 5.27 highlights the ratio of impaired to non-impaired staff. This can be used as an example base line if the actual value is unknown.

Table 5.27 shows that, within the UK, there is a large percentage of the working population that has a disability, as defined by the UK Government, and this calculation indicates that for every one officially categorised disabled person within a workforce, there are eight non-disabled members of the same workforce. Therefore this component highlights an example value which uses the above ratio of people with impairments. If a sample organisation of 20 is considered, then it would be expected that 2 members of that workforce has a impairment. This means that there are at least two users that might not be able to provide some of the common biometric samples and as mobility abnormalities are some of the most common impairments then techniques such as gait, facial or iris scans etc. may encounter problems [47].

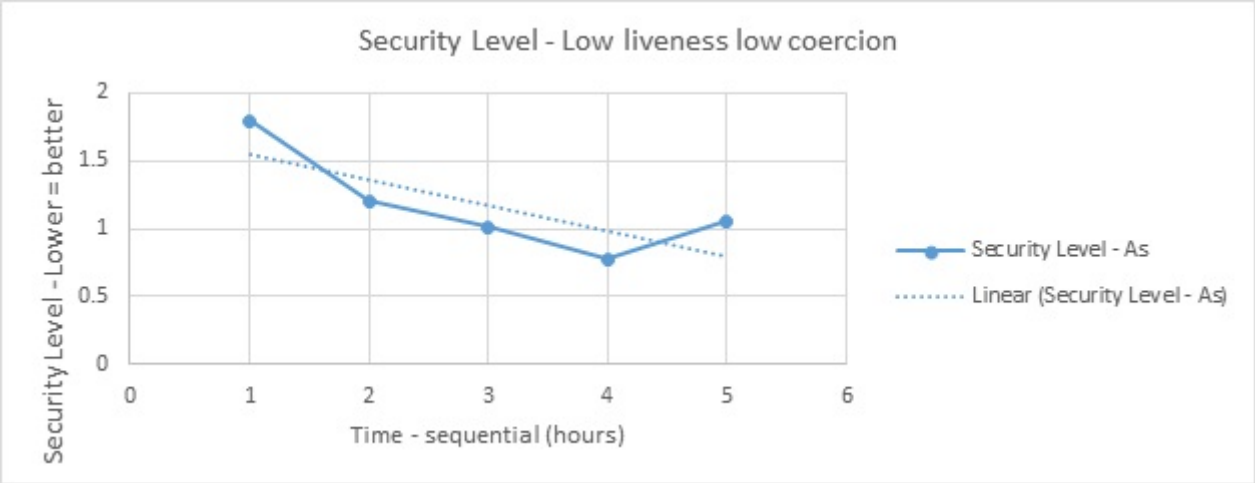


Fig. 5.9 Low liveness and low coercion.

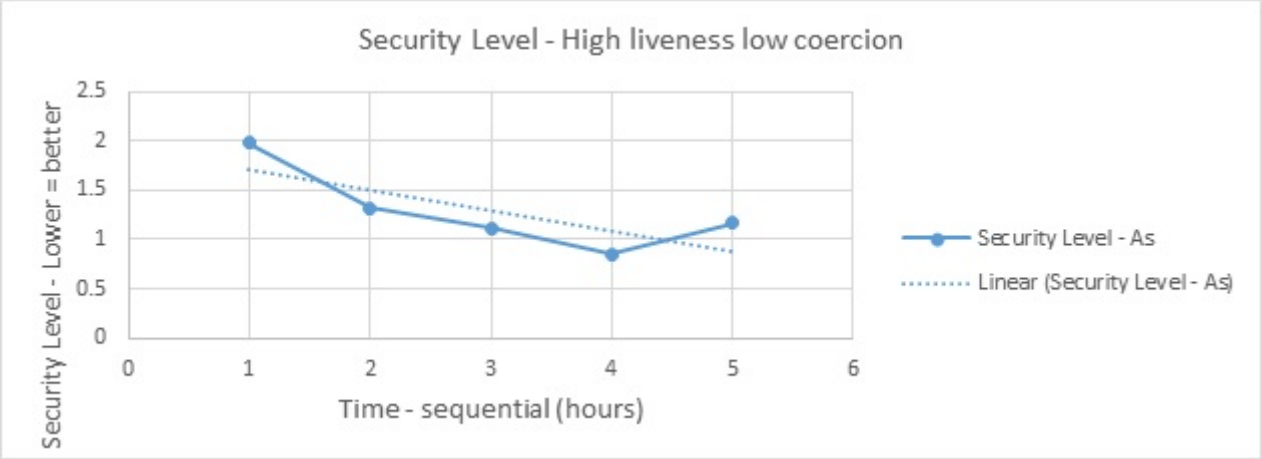


Fig. 5.10 High liveness and low coercion.

5.3 Taxonomy Evaluation

Whilst it has been shown that the taxonomy works there are certain issues to contend with. These factors demand further improvement and refinement and will make future iterations of the taxonomy and algorithm more robust. This section will cover what characteristics need improving and throughout the testing section six areas were highlighted. These are:

1. Interval to ratio scale
2. Data range clarification
3. Permanence refinement
4. Hardware reliance

Metric	Percent	Actual
Work-flow		
Actual		
Current Population	-	63,000,000
<14 Years	17.60%	11,088,000
>64 Years	16.40%	10,332,000
Non-workers	34.00%	21,420,000
Potential Working Population	66.00%	41,580,000
Population With Disability	16.40%	10,332,000
% of disabled population in workforce	46%	4,752,720
% of non-disabled population in workforce	76.00%	36,827,280
Ratio of disabled workforce to non-disabled workforce	7.748674	8:1

Table 5.27 Calculation of approximate impairment spread in workforce

5. Lack of mobile relevancy

6. Sample data and acceptance

Therefore the following section will consider the points raised and the best technique to solve, or alleviate the effects.

5.3.1 Interval to Ratio Scale

Currently, the scale being used has been interval that is limited to five points. This technique was chosen to limit the potential variations that could occur within each metric and whilst it could be argued that using a true ratio based system would be superior, as it would allow for extremes to be catered for, the more structured metric was instead implemented. This was done to allow current research to be more precisely fit into the taxonomy. Otherwise, it could have suffered from many different extremes of data potentially detracting from the efficiency aimed for. When the taxonomy has had more data applied to it then it will be at this point that the true use of interval will appear, as it will avoid a plethora of extreme data spikes instead ranging research more efficiently, and allowing a more practical use for researchers and developers. However as time progresses further iteration on the taxonomy will be needed to highlight if this is truly the case or whether a ration scale may indeed be more suitable.

5.3.2 Data Clarification

Data is often identified with a variety of metrics, units and standards, this is endemic throughout all of computing and biometric data is no different. Whilst there are some universal standards that are used there is also deviation which causes problems for this taxonomy. For example standards such as FAR/FRR/FTE/EER are commonly used throughout biometric device development, but a selection of researchers utilises their own standards and deviate from this general standardisation. For example within the research into facial spoof detection systems, [111] uses some of the traditional standards such as FAR and FRR, however the inclusion of a novel standard CIR (Correct Identification Rate) is also used. These standards are not used throughout other research and techniques therefore creating a higher degree of standard redundancy.

Within biometric research identification of the time taken to identify a sample is rarely identified and can range from less than 20 milliseconds to over 5 seconds. This means that comparing time is very difficult and needs to become more commonplace in biometric research. The inclusion of time in this taxonomy promotes its gathering and hopefully will promote the collection as a standard practice. if this was done than the taxonomy would improve as it would exclude more techniques.

A common accuracy and effectiveness measurement is also needed, as currently there are a number of factors that can represent the effectiveness of a biometric security and liveness techniques, but none for coercion. One of the main techniques is ERR and whilst it contains some negative attributes it is still a widely used standard that provides a baseline level. Therefore it can be used to compare different techniques in a robust manner. This is not to say that ERR should be used solely, however the development of a more universal measure would enable the maintainability of an overall standard therefore allowing an easy transition for new techniques to measure accuracy.

Spoof development has often be cited as the major flaw for biometric devices and this has become apparent when viewing the taxonomy data. This performance based classifier identifies the amount of techniques that can be used to spoof a specific sample, for example fingerprint scored very poorly due to the amount of ways to both gather and spoof a fingerprint e.g. gelatine, clay, photograph, wax, etc. [171] [105]. Whilst this is an important factor within this overall development of biometric techniques there is little to no standardisation and the taxonomy has highlighted this problem. Therefore there a standard must be developed that will identify the susceptibility to spoofing that can then be applied to the taxonomy. One such standard could be SSR – Spoof Susceptibility Rate: this would be able to identify the amount of spoofing methods there are for a specific technique, the greater the methods they higher the value. For example fingerprint may have a SSR of 4 denoting the amount of spoof

techniques applicable, whereas an ECG technique may have a SSR of 1 as there are very few techniques applicable to that biometric. This would allow their overall system to be more thoroughly integrated within the taxonomy and subsequently into more security systems.

All of these factors leads to two potential conclusions, the first being that the taxonomy has to become so broad that it encapsulates whatever potential factors a researcher can provide. This would make the standard in which the information appears almost irrelevant. There are a number of negative factors associated with this method. Firstly sometimes there is no information provided within the research so unless the research is brand new then there is the potential for data to be simply lacking and no matter how encompassing the taxonomy it cannot create data from nowhere. Secondly it is impractical to try and accommodate ever term, data type etc. within the taxonomy. Thirdly, if a technique uses a specific metric and then a second technique uses a different one then there is the potential that there will be confusion between the two.

This is why the second conclusion is favoured. By using the taxonomy developers and researchers could provide data that is specifically request by the taxonomy. The techniques would have relevant ranges identified which would require the researchers to follow. This would improve the ease of data integration and provided a more easily comparative research environment. This would allow a more straightforward comparison to occur at all levels, instead of the mixed and piecemeal comparison currently in operation. Therefore the solution to this would be to require researchers to utilise the taxonomy as a descriptive check-list as well as its other features, therefore, by providing the data requested the overall knowledge about the technique and others could be improved and could be more easily accessed by anyone using the taxonomy.

5.3.3 Hardware reliance

One factor that has become more and more apparent, as testing has progressed, is the taxonomy's reliance on hardware. There are occasions, such as within the collectability and performance metrics, that the major negative attributes are based on the additional technology factor. Whilst this is a legitimate factor care must be taken to make sure that the hardware inclusion does not overshadow other techniques. Within the current data collection range the overall impact is not too problematic however as more data is entered into the taxonomy this factor become more problematic. In fact the inclusion of hardware can be a very positive factor as it can provide a more solid basis of security, however in most cases the addition of more security threat vectors detracts from the positive factors brought in.

Therefore throughout the development of this taxonomy the impact of hardware has been considered and it has been noted that whilst currently a minor issue, it will only become more problematic when substantially more data has been included.

5.3.4 Lack of mobile relevancy

One factor that has been identified due to the work conducted by [152] is the effect of mobile system integration. Most full biometric systems involve stationary installations and it has only been recently that smart devices and techniques such as [152]'s have been become more prevalent. The mobile integration of biometrics still has a long way before regular inclusion occurs, and for the quality to be the same as static techniques. For this reason currently the taxonomy has a major emphasis towards traditional installations not making any specific allowances for mobile systems. This is something that may need to be considered in the future as a bolt on features however, for the current taxonomy, this is not within the scope.

5.3.5 Acceptance

One interesting factor that became apparent whilst developing the taxonomy was the importance of the acceptance classifier; as this will denote how easily the technique can be integrated into the overall system. Also how easily circumvention will affect the users understanding and acceptance of any techniques being suggested. This is an important attribute to consider and demands relevant research however it is something that is often ignored or glossed over within a research paper in favour of the more technical attributes of the technique. Whilst there should not be a wholesale change from technical attributes to this social factor: it should still be considered.

Within this taxonomy testing the problem occurred that there was little to no information regarding the acceptance of the individual techniques within the papers, and whilst there is research that identifies the acceptance of overall techniques the specifics are often ignored. Therefore it was almost impossible to adapt the data within the testing to denote the true acceptance. Therefore a decision was needed: should the data be kept at a standard level for all techniques therefore minimising the effect; or should the acceptance classifier be taken out completely removing the problem at its core.

The former point was the most suitable option for a number of reasons. Firstly if the acceptance classifier was taken out then it would change the overall effect of the taxonomy and would produce results that would not be as robust. This is because, as the acceptance section discussed above indicates, acceptance of a technique can lead to easier integration within systems. Therefore the removal of the classifier would cause irreparable harm to the

integrity of the overall taxonomy and whilst it would alleviate the problem of data validity the cost would be too great.

Secondly, the advantage of providing specific standard dummy data in this classifier allows it to exist but removes the immediate impact until all relevant data could be properly correlated and this could be done in one of two ways. The first is to utilise the clarification technique, identified in the Data Clarification section, promoting the use of the taxonomy in future research. This could be done by producing a number of research papers using the taxonomy to show its value in a practical setting, therefore promoting other researchers in their own understanding and acceptance of the taxonomy.

Therefore for the purpose of this testing it is decided to provide each technique with the same quantity of dummy data therefore providing the same amount of impact numerically. Whilst this is an adequate solution the main negative is that the data is not correct and only valid data will completely test the overall relevancy of the taxonomy, however due to the lack of available data this is not viable to immediately test, however it is one of the major areas that will be considered within the future work section.

These factors have been discerned from an in-depth analysis and evaluation of the taxonomy and the testing therein. It is obvious that the taxonomy works correctly and can provide interesting and relevant information to use by researchers and developers alike.

5.4 Algorithm Evaluation

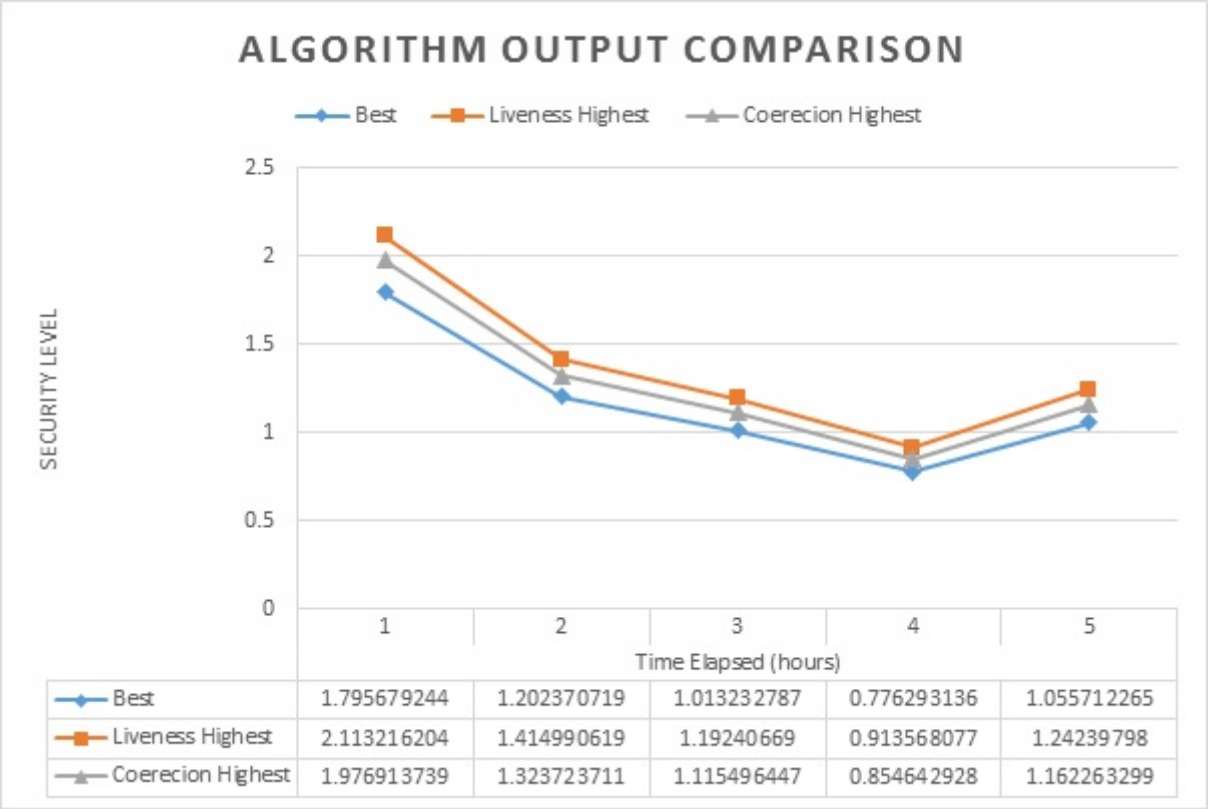
After the evaluation of the taxonomy the second area to consider is the evaluation of the algorithm. The algorithm has been populated with data, which is contained within Appendices Two – Raw Taxonomy and Algorithm Data. During this testing and analysis some factors became apparent that are of significant importance when utilising the overall algorithm within a developer/researcher area. These areas are:

1. Single Data Reliance
2. Data Specificity
3. Interface development

5.4.1 Single Data Reliance

The initial observation when using this algorithm is that whilst it can correctly identify the security level of the overall system, the focus is on one specific set of data. For example one authentication style over period of time. To take full advantage of this algorithm a method to

compare to other automatically and find the most suitable technique for a specific situation. Therefore if this comparison was identified the following would be seen, and then can be compared together, as shown by Figure 5.11.



Best	1.795679244	1.202370719	1.013232787	0.776293136	1.055712265
Liveness Highest	2.113216204	1.414990619	1.19240669	0.913568077	1.24239798
Coerecion Highest	1.976913739	1.323723711	1.115496447	0.854642928	1.162263299

Fig. 5.11 Real coercion and liveness data

When this comparison has been achieved it becomes easier to identify techniques that are superior to the others as they are numerically different from each other. In this case it is obvious that the "best" technique is the most secure, and the technique that had the highest coercion and lowest liveness was the second most secure and most practical. When viewing this data other factors need to be considered, as this is not intended to be the only degree of information, instead this is the top level component of analysis. Coupled with the data gathered from the taxonomy detailing which areas are providing the most impact, a developer/researcher is able to identify what factors have the most impact within a specific section alongside the overall algorithm providing a throughout tool that allows a more robust understanding of the overall system.

5.4.2 Data Specificity

The second factor that became apparent when dealing with the algorithm is that there is need for specific data and the more data it can parse the more robust and accurate it will become. This is because the algorithm uses real data to provide information about a technique or system over a period of time therefore demanding regular injections of data, which in the current environment is very difficult to provide. Therefore to improve the overall system more and more data needs to be injected into the algorithm in much the same way as the overall taxonomy. Coupled with this innate scalability, the practicality of including specific data, e.g. such as the data regarding the amount of users and login attempts throughout the system, is something that is very important to include as it provides specific information about the requirements. Within this algorithm a rough ratio of 1:8 has been used to detect if a user has any problems utilising biometric environment, however the actual data would be much more relevant to use within the overall algorithm.

5.4.3 Interface development

One factor that is not directly related to the technical areas of the algorithm, but has an enormous influence on the usefulness, is the lack of a user interface. The current algorithm is quite complicated to calculate however with the integration of a valid interface the algorithmic output could become much easier to work with and relevant to different user. This interface would enable researchers and developers to enter their data into easy to use forms which would convert the data, using the algorithm into meaningful outputs, including graphs etc. This is something that would be advantageous to the overall system, whilst not being urgent, therefore it would be suitable for future work to identify.

5.5 Conclusion

Throughout this testing, it has become apparent that the taxonomy has a lot of potentials and can be used to achieved a variety of factors. There are issues that have become apparent, such as the reliance on additional hardware and the ever-present importance of user acceptance in biometric security.

The techniques discussed in the liveness section area have all been chosen to show that the taxonomy works correctly. Despite some minor factors that need to be addressed, it has been proven that the taxonomy works as intended, and can provide relevant data to designers and researchers. Regardless this is only part of the taxonomies purpose; the secondary purpose is to check that it works within a coercion environment. However due to the comparative

originality of many coercion techniques it is much harder to gather data. Despite the problems surrounding this data gathering the taxonomy has worked for coercion and has provided similar outputs to the limited research that was considered.

While a variety of coercion techniques have been proposed and tested currently the skin conductivity response test and integration false authentication techniques hold the most promise, this is due to a combination of both current precedents within different areas of computing, and accuracy of the data. However it became clear that while, from an initial view, tangible key techniques could be viable, the testing indicated differently. It indicated that this technique was one of the poorest, mainly due to its ultimate reliance on additional hardware and the ease of access for an attacker.

The algorithm has used some different characteristics and has looked at the different components that can affect the security of a system. While this has produced positive results there are definitive issues to contend with, mainly focusing on the reliance on certain data types and the lack of user interface. Currently, the algorithm refinement sections were most successful as the enabled the reduction of unnecessary factors from the final algorithm, such as the segmenting of the liveness and coercion factors. This enables the overall algorithm to be much more robust by not having superfluous factors.

One thing that has been highlighted is that there is a need for much more data to get a better understanding of this taxonomy. While the data provided does indicate an initial concept, in the future, the inclusion of more data will help identify some of the flaws that have been identified such as the reliance on hardware, the use of interval instead of ratio system, etc.

Chapter 6

Conclusion

> Several times he had to flatten himself against the shelves as a thesaurus thundered by. He waited patiently as a herd of Critters crawled past, grazing on the contents of the choicer books and leaving behind them piles of small slim volumes of literary criticism
>
> Pratchett, 2004

The aims of this research have been to ...*categorise and evaluate different biometric liveness detection techniques alongside the development and classification of novel coercion detection techniques that will enable future techniques to be implemented.*

To achieve this: steps were identified using a grounded theory methodology. To begin, an in-depth analysis of current biometric literature had to be undertaken to highlight gaps in the current research field. The lack of classification, standardisation and comparative tools become the main thrust of the research. Methods of addressing this issue were considered, and the development of a taxonomy that would incorporate liveness techniques was a key goal. This was then extended also to be applicable for any future security techniques, as to do otherwise would mean the same problem regarding to much research to few comparative tools.

Therefore the goal of this research has been to create a taxonomy and algorithm that allows dynamic security environments to be developed using biometric systems and to expand on liveness and coercion detection. This is done by creating a comparative tool, the taxonomy, which can then be used to populate a system with the most viable security techniques in a participial area (the algorithm).

This research has highlighted a number of key areas that are very important to the overall system and to security generally. The reliance on liveness detection, by biometric security, is well documented and covered both in this work and in other research, however as liveness

detection has become a mainstream component coercion detection has been identified as the next stage to improve biometric security.

The testing has indicated that the taxonomy works however there are some considerations that must be looked at. Some of the most relevant factors indicated that one major flaw the system has is that it relies on the data validity of the taxonomy, and currently, for obvious reasons, there is comparatively little data available. Therefore in the future an important factor would be to get as much data parsed through the taxonomy, as the more data that is entered the higher the accuracy and validity and more understanding of the algorithm will occur.The other major factor that became apparent was the importance of user acceptance, whilst very important within the overall taxonomy, due to the lack of research in this area and the lack of ubiquity in classifiers, more data would be needed to validate this factor.

Overall the results show that the taxonomy works and can be useful in the development of dynamic systems by providing the data to the algorithm which also works correctly and can output degrees of security that can be used by the greater system. There are some potential issues that need addressing but none of which are severe enough to cause the system to fail, however to make the overall research more robust more work in the future needs to be conducted to highlight and fix any potential flaws or areas of inefficiency.

Finally the aims and objectives will be briefly discussed to identify were the research has succeeded or failed to achieve. The aim of the research we to *"identify the specific factors that liveness and coercion detection requires, coupled with the most suitable methods of implementation that can be derived from both"*. This has been done in a number of areas, the taxonomy has helped identify what coercion and liveness factors are most important to a biometric implementation. It also allows a user to identify the most suitable liveness or coercion standards to use within a system, allowing a more practical approach to technique fusion.

Whilst this is a very broad overview it does highlight how the research has achieved the aims of the research. To further understand this the objectives are also being considered, as this provides a more in-depth highlighting process.

1. "*Critically evaluate, in sufficient detail, the different algorithms, models, architectures and associated technologies within the biometric environment"*. This was done initial during the literature review chapter and highlighted the main factors with liveness detection. The emphasis on spoof sample effectiveness, the lack of standardisation and the impact of fusion between biometric device and liveness standards. This was subsequently mimicked within the coercion research as it highlighted the potential problems that coercion detection will face, such as the lack of research, the use of non-traditional techniques to gather data (IFA/TKTs etc.), along with issues surrounding

the acceptance of techniques by the users. Whilst these factors could have simply been classified they would not have provide the same depth of comparison that a taxonomy brings.

2. *"Develop a taxonomy for security techniques that can cover a range of different current and future technologies. Focusing on liveness detection, but maintaining scalability allowing future techniques to be adapted for the taxonomy whilst creating output that can be compared to other techniques, therefore creating a comparative and analytical tool"*. The taxonomy was developed to highlight the factors made within objective one. The aim was to make the taxonomy as scalable as possible enabling future technologies to be easily integrated into it. Therefore the decision was to utilise some of the pre defined categories that biometric device use as this would provide a familiar base level for future technology to adhere to. The main factor that was required was a way to easily use the taxonomy as a analytical tool, after identifying that the current ordinal method of measurement was very poor at doing this: it was decided to use an interval measuring system. This allows a valid and reliable comparison of techniques to be undertaken within uni-modal and multi-modal systems.

3. *"Investigate how the inclusion of different levels of liveness and coercion detection affects the security of the system"*. The development of the taxonomy was to provide the raw data that would allow comparison to occur. However on its own it would take a lot of work to manually compare all the techniques. Whilst this could be done, and is useful in certain circumstances such as identifying what factors cause a technique to score high or low. This highlighted the need for a process to easily compare techniques, ideally scenario specifically. Therefore an algorithm was considered to be the best method of highlighted the effects of different coercion and liveness standards. This algorithm would gather data from the scenario in question along with the liveness and coercion standards desired and would output a security level which could then be compared to find the most suitable fusion techniques for a system.

4. *"Identify the underlying structure of coercion detection and how it should work with biometric security, linking to both liveness detection and biometric authentication."* To achieve this comparison liveness detection was entered firstly, as this was the easier area to validate due to the availability of data. Coercion detection was harder to develop as the research concerning it is very limited. The overview of coercion detection was needed to highlight the salient factors it encountered such as noise, subtly and user acceptance.

5. *"Develop a baseline coercion detection environment, to determine the relevant validity of coercion detection in current biometric systems as well as how to measure coercion detection."*. After the basic environment was developed this objective intended to highlight how coercion detection could be developed. Numerous techniques were considered such as IFA, FMM etc. and whilst these techniques had advantages it became apparent that some were more subtitle than others. For example IFA was a high scoring technique than TKT.

6. *"Evaluate the taxonomy by testing it with a range of data gathered from respected peer reviewed research"*. The final objective was to wrap up all the previous research by evaluating the taxonomy and algorithm to identify the salient strengths and weakness as well as highlighting the suitability of the overall research. In numerous areas the taxonomy and algorithm performed satisfactory and allowed an interval based comparison to be used, however it also highlighted a number of problems that could lead to poor results or ineffectiveness such as the reliance on specific measurements.

Something that has become apparent throughout the process of this research is that there is no neat end point, the original aim and objectives have been achieved however the research has highlighted many areas that can be researched further. These different subjects and concepts have interesting and innovative features that could be applied to biometric security. Whilst the taxonomy has been created and the development of the architecture is developed there are still many areas that would be extremely interesting to consider, therefore whilst this research has come to its current endpoint the following section will identify where it could progress to in the future, alongside some of the associated areas that are applicable, as well as some of the current research outlets that can be applied for.

6.1 Future Work

Throughout this research there have been numerous occasions were a particular area become the focus for research, however it was not within the scope of the current research, or it was not directly linked but worthy of further commitment. This section highlights some of these areas indicating were the research could be taken, what aspects could be considered.

1. Context aware biometric subsystems. How context awareness can affect biometric devices.

2. Biometric residual data collection is a common problem, therefore identifying either materials that are most suitable for preventing the lifting of biometric samples.

3. Autonomous noise cancellation techniques would provide a huge bonus to security. However how would this be implemented, what are the parameters that would need to be taken into account?

4. Advance the taxonomy to include addition data from a larger variety of sources, both from official papers and practical primary data collecti?on techniques. This would be a direct progression from the PhD work.

5. As liveness detection techniques are normally more medical in nature than simple biometric authentication or verification techniques, what are the major factors that occur from this? What medical factors specifically effect the techniques, and what solutions can be developed to remove this problems.

6. Many of the different techniques that are identified as both liveness and coercion have very limited testing in the noise deviations that they face. A very applicable research would be to look into the concepts for example how cosmetics and topical medicines effect skin conductivity tests

7. Time based map of health using biometrics, medical collections from biometric ca provide a map of the user's health over a particular time. This could be aimed at specific time periods, or as a continuous detection method.

8. Leaving the technical features of this research for a moment, presuming that a user correctly identifies that they are being coerced what should the response be. Are there any legal or moral obligations from the system designers to do anything?

9. Many of the techniques utilise medical features to gather both liveness and coercion data, but how do medical anomalies effect his information. For example how does heart arrhythmia effect ECG results, irregular heartbeat, stroke etc.

10. Due to the very new area of coercion detection, how will fusion react when considering security, liveness and coercion? This is looked at briefly in the current research but more in-depth analysis would have opt be identified.

11. From a very practical standpoint, biometric security is becoming more and more poplar and utilised, however the research and information is still very much in an academic style, especially when detailing the mathematical imperatives surrounding the work. Therefore developing a dummies guide to biometric evaluation may be a valid research area. This could be linked to education as well by trying to find the best methods of teaching this kind of information.

12. Affect research is mainly based around skin conductivity response test, for coercion detection, what other data collection techniques could work.

6.2 Submission Goals

Throughout this research there has been numerous publications successfully submitted. However there are others that will be used as goals for future publications. This section highlights some of the main areas under consideration:

Name of Conference/Journal	Date and Location	Topic	Link
Science and Information Conference: Computing 2016 – London.	Submission 15th January (submitted)	Developing Coercion Detection Solutions for Biometric Security	http://saiconference.com/Computing2016 /CallforPapers
International Journal of Information Security	Submit Feb 19th 2016	Developing a Taxonomy for Liveness and Coercion Detection	http://www.springer.com/ computer/ security+and+cryptology /journal/10207
Computers and Security	Submission March – April 2016	Algorithm Development and Application	http://www.journals. elsevier.com/ computers-and-security
Information and Computer Security	Submission May-June 25th 2016	Algorithmic Development for Autonomous Biometric Security	//http://www.emerald group publishing.com/ics.htm

Table 6.1 Submission goals

References

[1] Mythbusters Episode 59 'Crimes and Myth-Demeanours 2', 2006.

[2] W Abernathy and L Tien. Biometrics: Who's Watching You?, 2003. URL http://www.eff.org/wp/biometrics-whos-watching-you.

[3] Aditya Abhyankar and Stephanie Schuckers. Fingerprint Liveliness Detection Using Local Ridge Frequencies And Multiresolution Texture Analysis Techniques. pages 1–4, Atlanta, 2006. IEEE International Conference on Image Processing. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4106531&tag=1.

[4] M. Abo-Zahhad, Sabah M. Ahmed, and S. N. Abbas. Biometric authentication based on PCG and ECG signals: Present status and future directions. *Signal, Image and Video Processing*, 8(4):739–751, 2014. ISSN 18631711. doi: 10.1007/s11760-013-0593-4. URL http://link.springer.com/10.1007/s11760-013-0593-4.

[5] Oasis Academy. Cashless Catering, 2014. URL http://www.oasisacademylordshill.org/content/cashless-catering.

[6] ACLM. ACLM Home Page, 2014. URL http://www.aclm.org.uk/.

[7] A Adler and S Shuckers. Overview of Biometric Vulnerabilities. In S Li and A K Jain, editors, *Encyclopedia of Biometrics*, pages 160–167. Springer, 2012.

[8] M Adolph. Biometric and Standards - ITU-T Technology Watch Report. Technical report, 2009. URL http://www.itu.int/dms_pub/itu-t/oth/23/01/T230100000D0002PDFE.pdf.

[9] S. Adolph, W. Hall, and P. Kruchten. Using grounded theory to study the experience of software development. *Empirical Software Engineering*, 16(4):487–513, January 2011. ISSN 13823256. doi: 10.1007/s10664-010-9152-6. URL http://link.springer.com/10.1007/s10664-010-9152-6.

[10] Foteini Agrafioti, Francis M. Bui, and Dimitrios Hatzinakos. Medical biometrics: The perils of ignoring time dependency. In *IEEE 3rd International Conference on Biometrics: Theory, Applications and Systems, BTAS 2009*. IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems, 2009. ISBN 9781424450206. doi: 10.1109/BTAS.2009.5339042.

[11] Av Aho, Je Hopcroft, and Jd Ullman. *The design and analysis of computer algorithms*. Addison-Wesley, 1987. ISBN 0-201-00029-6. URL http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+Design+and+Analysis+of+Computer+Algorithms#1.

[12] A Ali, F Deravi, and S Hoque. Directional Sensitivity of Gaze-Collinearity Features in Liveness Detection. *Emerging Security Technologie*, pages 8–11, 2013.

[13] G Allan. A critique of using grounded theory as a research method. *Electronic Journal of Business Research Methods*, 2(1):1–10, 2002.

[14] J Amaral. About Computing Science Research Methodology, 2011. URL http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.124.702.

[15] J. R. Angell. *Bodily Changes in Pain, Hunger, Fear and Rage; An Account of Recent Researches into the Function of Emotional Excitement*, volume 42. D. Appleton and Co., New York and London, 1915. ISBN 0843400781. doi: 10.1126/science.42.1089.696-a.

[16] P Art. Equality Act 2010, 2010. URL http://www.legislation.gov.uk/ukpga/2010/15/contents.

[17] C S Avilla, J G Casanova, F Ballesteros, L J M Garcia, M F A Gomez, D S Sierra, and G B Polzo. State of the art of mobile bioemtrics, liveness and on-coercion detection. Technical report, 2014.

[18] a Bandura. Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, 84(2):191–215, 1977. ISSN 0033-295X. doi: 10.1037/0033-295X.84.2.191.

[19] Wei Bao Wei Bao, Hong Li Hong Li, Nan Li Nan Li, and Wei Jiang Wei Jiang. A liveness detection method for face recognition based on optical flow field. In *2009 International Conference on Image Analysis and Signal Processing*. IASP 2009. International Conference In Image Analysis and Signal Processing, 2009. ISBN 978-1-4244-3987-4. doi: 10.1109/IASP.2009.5054589.

[20] N Bartlow and B Cukic. The vulnerabilities of biometric systems – an integrated look and old and new ideas'. Technical report, West Virginia University, West Virginia University, 2005.

[21] Nick Bartlow and Bojan Cukic. Evaluating the reliability of credential hardening through keystroke dynamics. *Proceedings - International Symposium on Software Reliability Engineering, ISSRE*, pages 117–126, 2006. ISSN 10719458. doi: 10.1109/ISSRE.2006.25.

[22] Houda Benaliouche and Mohamed Touahria. Comparative study of multimodal biometric recognition by fusion of iris and fingerprint. *The Scientific World Journal*, 2014:13, 2014. ISSN 1537744X. doi: 10.1155/2014/829369.

[23] Dm Berry, Mw Godfrey, Ric Holt, Cj Kapser, and I Ramos. Requirements Specifications and Recovered Architectures as Grounded Theories. *Groundedtheoryreview.Com*, 12(1):1–10, 2013. URL http://groundedtheoryreview.com/wp-content/uploads/2013/06/ARandRE.pdf.

[24] S. Bethune and J. Panlener. Stress a major health problem in the U.S., warns APA., 2007. URL http://www.apa.org/news/press/releases/2007/10/stress.aspx.

[25] B. Bhanu and V. Govindaraju. *Multibiometrics for Human Identification*. Cambridge University Press, New York, 2011. ISBN 9780511921056. doi: 10.1017/CBO9780511921056. URL http://ebooks.cambridge.org/ref/id/CBO9780511921056.

[26] L. Biel, O. Pettersson, L. Philipson, and P. Wide. ECG analysis: a new approach in human identification. *IMTC/99. Proceedings of the 16th IEEE Instrumentation and Measurement Technology Conference (Cat. No.99CH36309)*, 1(3):808–812, 1999. ISSN 1091-5281. doi: 10.1109/IMTC.1999.776813.

[27] Andrea Borghini. Nominalism and Realism - Philosophical Theory, 2015. URL http://philosophy.about.com/od/Philosophical-Theories-Ideas/a/Nominalism-And-Realism.htm.

[28] T. E. Boult, W. J. Scheirer, and R. Woodwork. Revocable fingerprint biotokens: Accuracy and security analysis. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007, 2007. Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on. ISBN 1424411807. doi: 10.1109/CVPR.2007.383110.

[29] Statistic Brain. Amputee Statistics, 2013. URL http://www.statisticbrain.com/amputee-statistics/.

[30] Mary Brandel. Biometrics: What, Where and Why, 2010. URL http://www.csoonline.com/article/2124928/identity-access/biometrics--what--where-and-why.html.

[31] P Brooks. Toolsmith. *ACM Comm*, pages 61–68, 1996.

[32] Bupa. Electrocardiogram (ECG), 2014. URL http://www.bupa.co.uk/individuals/health-information/directory/e/electrocardiogram.

[33] Bettina Burger, Dana Fuchs, Eli Sprecher, and Peter Itin. The immigration delay disease: Adermatoglyphia-inherited absence of epidermal ridges. *Journal of the American Academy of Dermatology*, 64(5):974–980, 2011. ISSN 01909622. doi: 10.1016/j.jaad.2009.11.013.

[34] J Busch and L Butcher. Getting Started with Business Taxonomy Design, 2007. URL https://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCIQFjAA&url=http%3A%2F%2Fwww.taxonomystrategies.com%2Fpresentations%2FGetting_Started.ppt&ei=-bDbU7bbFue20QWKx4HoCw&usg=AFQjCNFsoQA9zVRgHMHpGo9oL562Pu8ZCw&sig2=uNhUHf-L8l2SxInDgs3Qgg.

[35] J V Campellone. Parkinson's disease, 2013. URL http://www.nlm.nih.gov/medlineplus/ency/article/000755.htm.

[36] Vickie Chachere. Biometrics Used to Detect Criminals at Super Bowl, 2013. URL http://abcnews.go.com/Technology/story?id=98871&page=1#.UaEyOEDrzxc.

[37] Yao-Jen Chang Yao-Jen Chang, Wende Zhang Wende Zhang, and Tsuhan Chen Tsuhan Chen. Biometrics-based cryptographic key generation. In *2004 IEEE International Conference on Multimedia and Expo (ICME) (IEEE Cat. No.04TH8763)*, volume 3,

Taipei, 2004. Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '04). ISBN 0-7803-8603-5. doi: 10.1109/ICME.2004.1394707.

[38] G Chetty and M Wagner. Automated lip feature extraction for liveness verification in. *Proc. Image and Vision Computing*, pages 17–22, 2005.

[39] Heeseung Choi, Raechoong Kang, Kyoungtaek Choi, Andrew Teoh Beng Jin, and Jaihie Kim. Fake-fingerprint detection using multiple static features. *Optical Engineering*, 48(4):047202, 2009. ISSN 00913286. doi: 10.1117/1.3114606.

[40] Tanzeem Choudhury, Brian Clarkson, Tony Jebara, and Alex Pentland. Multimodal Person Recognition using Unconstrained Audio and Video. In *International Conference on Audio and Video-Based Person Authentication Proceedings*, pages 176–181, Washington DC, 1999. International Conference onAudio- and Video-Based Biometric Person Authentication.

[41] Nathan Clarke and Steven Furnell. Biometrics - The promise versus the practice. *Computer Fraud and Security*, 2005(9):12–16, 2005. ISSN 13613723. doi: 10.1016/S1361-3723(05)70253-0.

[42] Jeffrey F. Cohn. Foundations of human computing: Facial expression and emotion. In J A Coan and J B Allen, editors, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 4451 LNAI, pages 1–16. Oxford University Press, Oxford, 2007. ISBN 3540723463. doi: 10.1007/978-3-540-72348-6_1.

[43] Marco Conti, Sajal K Das, Chatschik Bisdikian, Mohan Kumar, Lionel M Ni, Andrea Passarella, George Roussos, Gerhard Tröster, Gene Tsudik, and Franco Zambonelli. Looking ahead in pervasive computing: Challenges and opportunities in the era of cyber–physical convergence. *Pervasive and Mobile Computing*, 8(1):2–21, 2012. URL http://dx.doi.org/10.1016/j.pmcj.2011.10.001$\delimiter"026E30F$npapers2://publication/doi/10.1016/j.pmcj.2011.10.001.

[44] Adam Czajka. Database of Iris Printouts and its Application : Development of Liveness Detection Method for Iris Recognition. In *MMAR, 18th International Conference*, pages 28–33, Miedzyzdroje, 2013. 18th International Conference on Methods and Models in Automation and Robotics (MMAR). ISBN 9781467355087.

[45] Ravi Das. Keystroke recognition. In Stan Z Li and A K Jain, editors, *Encyclopedia of Biometrics*, number 26, pages 29–31. Springer Reference, 2008.

[46] K. Delac and M. Grgic. A survey of biometric recognition methods. In *Proceedings. Elmar-2004. 46th International Symposium on Electronics in Marine*. 46th International Symposium In Electronics in Marine, 2004. ISBN 953-7044-02-5. doi: 10.1109/ELMAR.2004.1356372.

[47] Department for Work and Pensions. Disability prevalence estimates 2002/03 to 2011/12 (Apr to Mar) - Publications - GOV.UK, 2014. URL https://www.gov.uk/government/statistics/disability-prevalence-estimates-200203-to-201112-apr-to-mar.

[48] Dermalog. DERMALOG AFIS, 2013. URL http://www.dermalog.com/pdf/AFIS.pdf.

[49] I Dey. *Qualitative Data Analysis: A User-friendly Guide for Social Scientists.* Routledge, Oxon, 1993.

[50] Oxford English Dictionairies. Definition of coerce in English:, 2014. URL http://www.oxforddictionaries.com/definition/english/coerce.

[51] S D'Mello, A Graesser, and R Picard. Toward an Affect-Sensitive AutoTutor. *Intelligent Educational Systems*, pages 53–61, 2007.

[52] Michal Dolezel, Martin Drahansky, Jaroslav Urbanek, Eva Brezinova, and Tai-hoon Kim. Influence of Skin Diseases on Fingerprint Quality and Recognition. In Dr. Jucheng Yang (Ed.), editor, *New Trends and Developments in Biometrics*. 2012.

[53] R Edelberg and N R Burch. Skin resistance and galvanic skin response. Influence of surface variables, and methodological implications. *Archives of general psychiatry*, 7 (3):163–169, 1962. ISSN 0003-990X. doi: 10.1001/archpsyc.1962.01720030009002.

[54] Paul Ekman and Wallance V. Friesen. *The Facial Action Coding System.* Consulting Psycological Press, Palo Alto, 1978.

[55] Bmm El-Basioni, Sma El-kader, and Mahmoud Abdelmonim. Smart home design using wireless sensor network and biometric technologies. *Information Technology*, 2(3): 413–429, 2013. URL http://www.ijaiem.org/Volume2Issue3/IJAIEM-2013-03-29-098.pdf.

[56] Wafa Elgarah and Natalia Falaleeva. Adoption of Biometric Technology: Information Privacy in TAM. In *AMCIS 2005 Proceedings*, page Paper 222. AMCIS 2005 Proceedings., 2005. ISBN 9781604235531. URL http://aisel.aisnet.org/amcis2005/222.

[57] Malte Elson and Christopher J. Ferguson. Twenty-Five Years of Research on Violence in Digital Games and Aggression. *European Psychologist*, 19(1):33–46, January 2014. ISSN 1016-9040. doi: 10.1027/1016-9040/a000147. URL http://econtent.hogrefe.com/doi/abs/10.1027/1016-9040/a000147?journalCode=epp.

[58] Joseph Etherton, Marci Lawson, and Reiko Graham. Individual and gender differences in subjective and objective indices of pain: Gender, fear of pain, pain catastrophizing and cardiovascular reactivity. *Applied Psychophysiology Biofeedback*, 39(2):89–97, 2014. ISSN 10900586. doi: 10.1007/s10484-014-9245-x.

[59] Duquense Ethnography. *A critique of an ethnographic approach to the study of an online health support community: Advantages, disadvantages, and lessons learned*. Duquesne Ethnography, 2003.

[60] Leandro O. Freitas, Giovani R. Librelotto, Henrique G G Pereira, Jeferson Kasper, Ricardo G. Martini, Bruno Mozzaquatro, and Rafael T. Pereira. Applying pervasive computing in an architecture for homecare environments. In *Proceedings - IEEE 9th International Conference on Ubiquitous Intelligence and Computing and IEEE 9th International Conference on Autonomic and Trusted Computing, UIC-ATC 2012*, pages 685–692. 9th International Conference on Ubiquitous Intelligence and Computing

and 9th International Conference on Autonomic and Trusted Computing, 2012. ISBN 978-1-4673-3084-8. doi: 10.1109/UIC-ATC.2012.161.

[61] Steven Furnell and Konstantinos Evangelatos. Public awareness and perceptions of biometrics. *Computer Fraud and Security*, 2007(1):8–13, 2007. ISSN 13613723. doi: 10.1016/S1361-3723(07)70006-4.

[62] L Ghiani, P Denti, and G L Marchialis. Experimental results on fingerprint liveness detection. In *Articulated Motion and Deformable Objects*, pages 210–218. Springer Berlin Heidelberg, Berlin, 2012.

[63] Barney L Glaser and Anselm L Strauss. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. 1967. URL http://www.amazon.co.uk/The-Discovery-Grounded-Theory-Qualitative/dp/0202302601.

[64] T González, J Diaz-Herrera, and A Tucker, editors. *Computing Handbook, Third Edition: Computer Science and Software Engineering*. Chapman and Hall, 3rd edition. URL http://www.amazon.co.uk/Computing-Handbook-Third-Edition-Engineering/dp/1439898529.

[65] Lawrence a Gordon, Martin P Loeb, William Lucyshyn, and Robert Richardson. *Computer Crime and Security Survey*. Number 11. 2006. ISBN 11.

[66] Gov.uk. Work related stress - Research and statistics, 2014. URL http://www.hse.gov.uk/stress/research.htm.

[67] International Biometric Group. How is 'Biometrics' Defined?, 2007. URL http://www.biometricgroup.com/reports/public/reports/biometric_definition.html[accessed15/05/09].

[68] Payas Gupta and Debin Gao. Fighting Coercion Attacks in Key Generation using Skin Conductance. CA, 2008. USENIX Security'10 Proceedings of the 19th USENIX conference on Security.

[69] John Harp and Andrew Sayer. *Method in Social Science: A Realist Approach.*, volume 15. Routledge, Oxon, 1986. ISBN 0415076072. doi: 10.2307/2070081.

[70] J. Healey and R. Picard. SmartCar: detecting driver stress. In *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, volume 4, Barcelona, 2000. 2000. Proceedings. 15th International Conference on Pattern Recognition. ISBN 0-7695-0750-6. doi: 10.1109/ICPR.2000.902898.

[71] Improving Health and Care Worldwide. Cause and Effect Diagram, 2014. URL http://www.mmm.ethz.ch/dok01/d0000538.pdf.

[72] E Hedman, O Widler-Smith, M S Goodwin, M Poh, R Flectcher, and R Picard. iCalm: measuring electrodermal. N.J, 2009. 3rd International Conference on.

[73] R Hicks. Hyponatraemia, 2012. URL http://www.webmd.boots.com/a-to-z-guides/hyponatraemia.

[74] Hitachi. Planet Cash: Europe's first biometric ATM shared network, 2014. URL http://www.hitachi.eu/about/press/pdfs/Press_Release_Hitachi_ITC_14May2014FINALr.pdf.

[75] E Horowitz. Design and Classification of Algorithms. In *Encyclopedia Of Computer Science*, pages 33–37. 2003.

[76] John A Hughes and W.W Sharrock. *The Philosophy of Social Research John*. 1997. URL http://www.amazon.co.uk/Philosophy-Social-Research-Longman-Series/dp/0582311055.

[77] K Imaizumi. Treacher Collins syndrome, 1996. URL http://ghr.nlm.nih.gov/condition/treacher-collins-syndrome.

[78] C E Izard. *The maximally discriminative affect coding system (MAX)*. University of Delaware, Instructional Resource Center, Newark, 1979.

[79] A Jain, L Hong, and S Pankanti. An Identity-Authentication System. In *Proceding of the IEEE*, 1997.

[80] a K Jain and a Ross. U Uludag, Biometric template security: Challenges and solutions. In *Proc of European Signal Processing Conf (EUSIPCO)*, 2005.

[81] A K Jain, S Pankanti, S Prabhakar, L Hong, J L Wayman, and A Hong. Biometrics: A Grand Challenge. In *Proc. International Conference on Pattern Recognition (ICPR) VOL ii*, pages 935–942, Cambridge, 2004.

[82] Anil K Jain, Arun Ross, and Salil Prabhakar. An Introduction to Biometric Recognition 1. *IEEE Transactions on Circuits and Systems fir Vudei Technology*, 14(1):1–29, 2004.

[83] T James, T Pirim, and K Boswell. Determining the intention to use biometric devices: An application and extension of the technology acceptance model. *Journal of Organizational and End User Computing*, 18(3):1–24, 2006. ISSN 1546-2234. URL http://www.igi-global.com/article/journal-organizational-end-user-computing/3812.

[84] Abbas Javadtalab, Laith Abbadi, Mona Omidyeganeh, Shervin Shirmohammadi, Carlisle M. Adams, and Abdulmotaleb El Saddik. Transparent non-intrusive multimodal biometric system for video conference using the fusion of face and ear recognition. In *2011 9th Annual International Conference on Privacy, Security and Trust, PST 2011*, pages 87–92. Ninth Annual Conference on Privacy, Security and Trust, PST 2011,, 2011. ISBN 9781457705847. doi: 10.1109/PST.2011.5971968.

[85] Magne Jø rgensen, Tore Dybå, Knut Liestø l, and Dag I.K. Sjø berg. Incorrect results in software engineering experiments: How to improve research practices. *Journal of Systems and Software*, March 2015. ISSN 01641212. doi: 10.1016/j.jss.2015.03.065. URL http://linkinghub.elsevier.com/retrieve/pii/S0164121215000679.

[86] Peter Johnson and Richard Lazarick. Biometric Liveness Detection : Framework and Metrics - Day 4. In ... *Biometric* ..., 2012. URL http://biometrics.nist.gov/cs_links/ibpc2012/presentations/Day4/403_schuckers.pdf.

[87] Peter Johnson and Stephanie Schuckers. Evaluation of Presentation Attack Detection : An Example, 2014.

[88] O Kahm and N Damer. 2D face liveness detection: An overview. In *... (BIOSIG), 2012 BIOSIG-Proceedings of the ...*, Darmstadt, 2012. Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG), 2012 BIOSIG -. ISBN 9783885792901. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6313547.

[89] Masashi Kanematsu, Hironobu Takano, and Kiyomi Nakamura. Highly reliable liveness detection method for iris recognition. In *Proceedings of the SICE Annual Conference*, pages 361–364, Takamatsu, 2007. 2007 Annual Conference SICE. ISBN 4907764286. doi: 10.1109/SICE.2007.4421008.

[90] Geoffrey Keppel. *Statistics: A First Course. 3rd ed.*, volume 26. McGraw-Hill, Inc., 1981. doi: 10.1037/020575.

[91] Gahyun Kim, Sungmin Eum, Jae Kyu Suhr, Dong Ik Kim, Kang Ryoung Park, and Jaihie Kim. Face liveness detection based on texture and frequency analyses. In *Proceedings - 2012 5th IAPR International Conference on Biometrics, ICB 2012*, pages 67–72. 5th IAPR International Conference Biometrics (ICB),, 2012. ISBN 9781467303941. doi: 10.1109/ICB.2012.6199760.

[92] Sooyeon Kim, Sunjin Yu, Kwangtaek Kim, Yuseok Ban, and Sangyoun Lee. Face liveness detection using variable focusing. In *Proceedings - 2013 International Conference on Biometrics, ICB 2013*. lnternational Conference on Biometrics (ICB), 2013. ISBN 978-1-4799-0310-8. doi: 10.1109/ICB.2013.6613002.

[93] Sylvia D. Kreibig. Autonomic nervous system activity in emotion: A review, 2010. ISSN 03010511. URL http://homes.di.unimi.it/~boccignone/GiuseppeBoccignone_webpage/CompAff2012_files/Autonomicnervoussystemactivityinemotion-Areview.pdf.

[94] S Kung, M Mak, and S Lin. *Biometric Authentication: A Machine Learning Approach*. Prentice Hall Information and System Sciences Series, 2004.

[95] J Laird. SOAR: An architecture for general intelligence. *Artificial Intelligence*, 33(1): 1–64, September 1987. ISSN 00043702. doi: 10.1016/0004-3702(87)90050-6. URL http://www.sciencedirect.com/science/article/pii/0004370287900506.

[96] Ec Lee, Kr Park, and Jaihie Kim. Fake iris detection by using purkinje image. In D Zhang and A K Jain, editors, *Advances in Biometrics*, pages 397–403. Springer, 2005. ISBN 3540311114. URL http://link.springer.com/chapter/10.1007/11608288_53.

[97] Eui Chul Lee, You Jin Ko, and Kang Ryoung Park. Fake iris detection method using Purkinje images based on gaze position. *Optical Engineering*, 47(6):067204, 2008. ISSN 00913286. doi: 10.1117/1.2947582.

[98] Sz Z Li and Ak K Jain. *Handbook of Face Recognition*, volume 54. Springer, New York, 2005. ISBN 9780857299314. doi: 10.1007/978-0-85729-932-1. URL http://link.springer.com/content/pdf/10.1007/978-0-85729-932-1.

pdf$\delimiter"026E30F$nhttp://www.springerlink.com/index/10.1007/978-0-85729-932-1.

[99] Perttu J Lindsberg. Editorial comment–high blood pressure after acute cerebrovascular occlusion: risk or risk marker? *Stroke; a journal of cerebral circulation*, 36(2): 268–269, 2005. ISSN 1524-4628. doi: 10.1161/01.STR.0000153045.33710.bc.

[100] Simon Liu and Mark Silverman. Practical guide to biometric security technology. *IT Professional*, 3(1):27–32, 2001. ISSN 15209202. doi: 10.1109/6294.899930.

[101] André Lourenço, Hugo Silva, and Ana Fred. Unveiling the biometric potential of finger-based ECG signals. *Computational Intelligence and Neuroscience*, 2011:8, 2011. ISSN 16875265. doi: 10.1155/2011/720971.

[102] Wanli Ma, John Campbell, Dat Tran, and Dale Kleeman. Password entropy and password quality. In *Proceedings - 2010 4th International Conference on Network and System Security, NSS 2010*, pages 583–587, Melbourne, 2010. 2010 4th International Conference on Network and System Security (NSS). ISBN 9780769541594. doi: 10.1109/NSS.2010.18.

[103] Sung Won Park Marios Savvides, Jingu Heo. *Handbook of Biometrics*. 2008. ISBN 978-0-387-71040-2. doi: 10.1007/978-0-387-71041-9_10.

[104] Pablo Martinez-Lozano Sinues, Malcolm Kohler, and Renato Zenobi. Human Breath Analysis May Support the Existence of Individual Metabolic Phenotypes. *PLoS ONE*, 8(4), 2013. ISSN 19326203. doi: 10.1371/journal.pone.0059909.

[105] T Matsumoto, H Matsumoto, K Yamada, and S Hoshino. Optical security and counterfeit deterrence techniques IV. In *Proceedings of SPIE vol. #4677*, Japan, 2002.

[106] Tsutomu Matsumoto, Hiroyuki Matsumoto, Koji Yamada, and Satoshi Hoshino. Impact of Artificial “Gummy" Fingers on Fingerprint Systems. In *Proceedings of SPIE*, volume 4677, pages 275–289, 2002. ISBN 0819444170. doi: 10.1117/12.462719. URL http://cryptome.org/gummy.htm.

[107] V Matyás and Z Riha. Bioemtric authetnication - Security and usability. Proc. 6th IFIP TC6/TC11 Conf. Commun. Multimedia Security, 2002.

[108] MMU. Analyse This!!! Learning to analyse quantitative data, 2007. URL http://www.learnhigher.ac.uk/analysethis/main/quantitative1.html.

[109] Maxine Most. Acuity Market Intelligence - The Global Biometrics and Mobility Report: The Convergence of Commerce and Privacy, 2014. URL http://www.acuity-mi.com/GBMR_Report.php.

[110] Makram Nabti and Ahmed Bouridane. An effective and fast iris recognition system based on a combined multiscale feature extraction technique. *Pattern Recognition*, 41 (3):868–879, 2008. ISSN 00313203. doi: 10.1016/j.patcog.2007.06.030.

[111] B G Nalinakshi, S Hatture, M Gabasavalgi, and R Karchi. Liveness Detection Technique for Prevention of Spoof Attack in Face Recognition System. *International Journal of Emerging Technology and Advanced Engineering*, 3(12), 2013.

[112] B G Nalinakshi, S.Gabasavalgi Sanjeevakumar, M Hatture Manjunath, and Rashmi P Karchi. Liveness Detection Technique for Prevention of Spoof Attack in Face Recognition System. *International Journal of Emerging Technology and Advanced Engineering*, 3(12):627–633, 2013. URL http://www.ijetae.com/files/Volume3Issue12/IJETAE_1213_111.pdf.

[113] K Nandakumar. Security Issues, System Design. In S Li and A K Jain, editors, *Encylopedia of Biometrics*, pages 1152–1158. Springer, 2012.

[114] K Nandakumar, Y Chen, S Dass, and A Jain. No Title. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 30(2), 2008.

[115] Karthik Nandakumar, Student Member, Yi Chen, Sarat C Dass, and Anil K Jain. Likelihood Ratio Based Biometric Score Fusion. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 30(2):1–9, 2007.

[116] Karthik Nandakumar, Anil K. Jain, and Abhishek Nagar. Biometric template security. *Eurasip Journal on Advances in Signal Processing*, 2008, 2008. ISSN 16876172. doi: 10.1155/2008/579416.

[117] NHS Choices. How do I check my pulse? (Health questions), 2013. URL http://www.nhs.uk/chq/Pages/2024.aspx?CategoryID=52.

[118] Shailendra Nigam and Deepak Garg. Choosing Best Algorithm Design Strategy For A Particular Problem 2 . Comparison of algorithm Design Strategy 3 . Best Algorithm Design Selection for a particular problem. In *IEEE International Advance Computing Conference (IACC 2009)*, number March, pages 6–7, Patiali, 2009.

[119] T Ojala, M Pietikäinen, and T Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. In *Tpami*, volume 24, pages 971–987. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002.

[120] Leala Padmanabhan. Biometrics in smartphones need more control - ex-GCHQ boss - BBC News, 2014. URL http://www.bbc.co.uk/news/uk-politics-30211238.

[121] Gang Pan, Zhaohui Wu, and Lin Sun. Liveness Detection for Face Recognition. In K Delac, M Grgic, and M Steward Bartlett, editors, *Recent Advances in Face Recognition*, number December, page 236. I-Tech, Vienna, 2008. ISBN 9789537619343. URL http://www.intechopen.com/.

[122] R Picard. Human–Computer Coupling. *PROCEEDINGS OF THE IEEE*, 88(8), 1998.

[123] Rosalind W. Picard, Elias Vyzas, and Jennifer Healey. Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10):1175–1191, 2001. ISSN 01628828. doi: 10.1109/34.954607.

[124] P Pincher. A guide to developing taxonomies for effective data management, 2013. URL http://www.computerweekly.com/feature/A-guide-to-developing-taxonomies-for-effective-data-management.

[125] Wendi Pohs. Building a Taxonomy for Auto-classification. *Bulletin of the American Society for Information Science and Technology*, 39(2):34–38, 2013. ISSN 1931-6550. doi: 10.1002/bult.2013.1720390210.

[126] B Prasanalakshmi and A Kannammal. A Secure Cryptosystem From Palm Vein Biometrics. Seoul, 2009. ICIS.

[127] J Preece, Y Rogers, and H Sharp. *Interaction Design: Beyond human-computer interaction*. Wiley, 2002.

[128] Psychlab. SKIN CONDUCTANCE EXPLAINED, 2014. URL http://www.psychlab.com/SC_explained.html.

[129] M K Qureshi. Liveness detcion of biomtric traits. *International Journal of Information Technology and Knowledge Management*, 4(1):293–295, 2011.

[130] N. K. Ratha, J. H. Connell, and R. M. Bolle. Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal*, 40(3):614–634, 2001. ISSN 0018-8670. doi: 10.1147/sj.403.0614.

[131] Nalini K. Ratha, Sharat Chikkerur, Jonathan H. Connell, and Ruud M. Bolle. Generating cancelable fingerprint templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):561–572, 2007. ISSN 01628828. doi: 10.1109/TPAMI.2007.1004.

[132] S Rathus. *Psychology: Concepts and Connections*. Wadsmorth, Belmont, 2012.

[133] P. Venkata Reddy, Ajay Kumar, S. M K Rahman, and Tanvir Singh Mundra. A new antispoofing approach for biometric devices. *IEEE Transactions on Biomedical Circuits and Systems*, 2(4):328–337, 2008. ISSN 19324545. doi: 10.1109/TBCAS.2008.2003432.

[134] RESO. Edge Hill - Framework for Research Ethics, 2015. URL https://www.edgehill.ac.uk/governance/files/2013/03/Framework-for-Research-Ethics.pdf.

[135] Chris Roberts. Biometric attack vectors and defences. *Computers and Security*, 26(1):14–25, 2007. ISSN 01674048. doi: 10.1016/j.cose.2006.12.008.

[136] Kathryn a Roberts and Richard W Wilson. FORUM : QUALITATIVE SOCIAL RESEARCH ICT and the Research Process : Issues Around the Compatibility of Technology with Qualitative Data Analysis 2 . The Philosophy of Qualitative Data. *Qualitative Social Research*, 3(2), 2002.

[137] Roberto Roizenblatt, Paulo Schor, Fabio Dante, Jaime Roizenblatt, and Rubens Belfort. Iris recognition as a biometric method after cataract surgery. *Biomedical engineering online*, 3(2):2, 2004. ISSN 00029394. doi: 10.1186/1475-925X-3-2.

[138] Peter Rosenbaum and Debra Stewart. The World Health Organization International Classification of Functioning, Disability, and Health: A Model to Guide Clinical Thinking, Practice and Research in the Field of Cerebral Palsy. *Seminars in Pediatric Neurology*, 11(1):5–10, 2004. ISSN 10719091. doi: 10.1016/j.spen.2004.01.002.

[139] A Ross. Multibiometrics. In *Encylopedia of Biometrics*, pages 967–973. Springer, 2012.

[140] A Ross and A Jain. Multimodal biometrics: An overview. 1. In *2th European Signal Processing Conference (EUSIPCO)*, pages 1122–1221, 2004.

[141] Gordon Rugg and Marian Petre. *Gentle Guide to Research Methods*. McGraw-Hill International, 2006. ISBN 9780335230198. URL http://site.ebrary.com/lib/uoh/docDetail.action?docID=10197031.

[142] Per Runeson and Martin Höst. Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering*, 14(2):131–164, December 2009. ISSN 13823256. doi: 10.1007/s10664-008-9102-8. URL http://link.springer.com/10.1007/s10664-008-9102-8.

[143] K Jain Salil, Prabhakar and Sharath, Pankanti and Anil. Biometric recognition: Security and privacy concerns. *Security & Privacy, IEEE*, 1(2):33–42, 2003.

[144] Marie Sandström. Liveness detection in fingerprint recognition systems. Technical report, Sweden, 2004. URL $\delimiter"026E30F$nhttp://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-2397.

[145] Maggi Savin-Baden. *Problem-based Learning in Higher Education : Untold Stories*, volume 21. 2000. ISBN 0335203388. doi: 10.1080/07294360220124620. URL http://mhc.mcgraw-hill.co.uk/openup/chapters/033520337X.pdf.

[146] Michael a. Sayette, Jeffrey F. Cohn, Joan M. Wertz, Michael a. Perrott, and Dominic J. Parrott. A psychometric evaluation of the facial action coding system for assessing spontaneous expression. *Journal of Nonverbal Behavior*, 25(3):167–185, 2001. ISSN 01915886. doi: 10.1023/A:1010671109788.

[147] Markus Schatten, Miroslav Bača, and Kornelije Rabuzin. A taxonomy of biometric methods. In *Proceedings of the International Conference on Information Technology Interfaces, ITI*, pages 389–393. 30th International Conference on Information Technology Interfaces, 2008. ISBN 9789537138127. doi: 10.1109/ITI.2008.4588441.

[148] Markus Schatten, Miroslav Baca, and Mirko Cubrilo. Towards a General Definition of Biometric Systems. *IJCSI International Journal of Computer Science Issues*, 2, 2009. URL http://cogprints.org/6690/.

[149] Stephanie Schuckers and L Hornak. Issues for liveness detection in biometrics, 2002. URL http://www.biometrics.org/bc2002/2_bc0130_DerakhshabiBrief.pdf.

[150] Jochen Schwarze. *Introduction to the design and analysis of algorithms*, volume 4. Villanove University, 1980. ISBN 0321364139. doi: 10.1016/0377-2217(80)90009-0.

[151] Graeme Shanks. Guidelines for conducting positivist case study research in information systems. *Australasian Journal of Information Systems*, 10(1):76–85, November 2002. ISSN 13262238. doi: 10.3127/ajis.v10i1.448. URL http://journal.acs.org.au/index.php/ajis/article/view/448.

[152] Khairul Azami Sidek, Vu Mai, and Ibrahim Khalil. Data mining in mobile ECG based biometric identification. *Journal of Network and Computer Applications*, 44: 83–91, September 2014. ISSN 10848045. doi: 10.1016/j.jnca.2014.04.008. URL http://www.sciencedirect.com/science/article/pii/S1084804514000915.

[153] Khairul Azami Sidek, Vu Mai, and Ibrahim Khalil. Data mining in mobile ECG based biometric identification. *Journal of Network and Computer Applications*, 44:83–91, 2014. ISSN 10958592. doi: 10.1016/j.jnca.2014.04.008.

[154] Terence Sim and Rajkumar Janakiraman. Are digraphs good for free-text keystroke dynamics? In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Singapore, 2007. In proceeding of: Computer Vision and Pattern Recognition. ISBN 1424411807. doi: 10.1109/CVPR.2007.383393.

[155] Yogendra Narain Singh and Sanjay Kumar Singh. A taxonomy of biometric system vulnerabilities and defences. *International Journal of Biometrics*, 5(2):137, 2013. ISSN 1755-8301. doi: 10.1504/IJBM.2013.052964. URL http://www.scopus.com/inward/record.url?eid=2-s2.0-84877275048&partnerID=tZOtx3y1.

[156] W Stallings and L Brown. *Computer Security: Principles and Practice*. Prentice Hall, 2012.

[157] Hanne Storm. Changes in skin conductance as a tool to monitor nociceptive stimulation and pain. *Current opinion in anaesthesiology*, 21(6):796–804, 2008. ISSN 0952-7907. doi: 10.1097/ACO.0b013e3283183fe4.

[158] M Sung and A Pentland. PokerMetrics: Stress and Lie Detection through Non-Invasive Physiological Sensing, 2005. URL http://citeseer.uark.edu:8080/citeseerx/showcitingjsessionid=5DB10DCA426C12D7434B594288FD67E8?doi=10.1.1.153.9203&sort=ascdate.

[159] Wen-pei Sung. *Computer, Intelligent Computing and Education Technology*, volume 1. CRC Press, 2014. ISBN 978-1-138-02469-4. doi: 10.1201/b16698. URL http://www.crcnetbase.com/doi/book/10.1201/b16698.

[160] Katharina Tabbert, Rudolf Stark, Peter Kirsch, and Dieter Vaitl. Dissociation of neural responses and skin conductance reactions during fear conditioning with and without awareness of stimulus contingencies. *NeuroImage*, 32(2):761–770, 2006. ISSN 10538119. doi: 10.1016/j.neuroimage.2006.03.038.

[161] C Tapper. CSI computer crime and security survey. Technical Report Pt 3, 2008.

[162] Lisa Thalheim, Jan Krissler, and Peter-Michael Ziegler. Body check: biometric access protection devices and their programs put to the test, 2002. URL http://www.larc.usp.br/~pbarreto/Leitura2-Biometria.pdf.

[163] L Thorogood. Performance evaluation., 2009. URL http://www.csee.wvu.edu/~natalias/biom426/performance_fall09.pdf.

[164] Ying-Li Tian Ying-Li Tian, T. Kanada, and J.F. Cohn. Recognizing upper face action units for facial expression analysis. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*, 1(2):97–115, 2000. ISSN 1063-6919. doi: 10.1109/CVPR.2000.855832.

[165] Walter F. Tichy. Should computer scientists experiment more?, 1998. ISSN 00189162. URL http://www.cs.ou.edu/~fagg/classes/empirical_methods_2006/papers/ShouldComputerScientistsExperimentMore.pdf.

[166] Bori Toth. Biometrics Biometric Liveness Detection. *Information Security Bulletin*, 10(October):291–298, 2005.

[167] Gracian Trivino, Luis Mengual, and Albert van der Heide. Towards an architecture for semiautonomous robot telecontrol systems. *Information Sciences*, 179(23):3973–3984, 2009. ISSN 00200255. doi: 10.1016/j.ins.2009.08.007.

[168] a Uhl and P Wild. Experimental evidence of ageing in hand biometrics. *Biometrics Special Interest Group (BIOSIG), 2013 International Conference of the*, 2014:1–6, 2013. ISSN 16175468.

[169] Umut Uludag and Anil K. Jain. Attacks on Biometric Systems: A Case Study in Fingerprints. In *Proceedings of SPIE*, volume 5306, pages 622–633. In Proceedings of SPIE (Vol. 5306, pp. 622-633)., 2004. doi: 10.1117/12.530907. URL http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=837770.

[170] Clarkson University. Clarkson University Engineer Outwits High-Tech Fingerprint Fraud', 2005. URL http://www.sciencedaily.com/releases/2005/12/051216193022.htm[accessed15/07/09].

[171] T. van der Putte and J. Keuning. *Biometrical fingerprint recognition: don't get your fingers burned*, volume 31. Kluwer Academic Publisher, 2000. ISBN 0-7923-7953-5. URL http://books.google.com/books?hl=en&lr=&id=mGOnonNnr7AC&oi=fnd&pg=PA289&dq=Biometrical+fingerprint+recognition:+don't+get+your+fingers+burned&ots=8azqbsbWZH&sig=yQMdezyU5gQ6bN4Kg4hPZ7p6yrQ.

[172] O Vermesan. Outlook on Future IoT Applications. In H Sundmaeker, P Guillemin, P Friess, and S Woelfflé, editors, *Vision and Challenges*, pages 181–190. European Commission Information Society and Media, 2010.

[173] Heinrich Wansing. Doxastic Decisions, Epistemic Justification, and The Logic of Agency. *Philosophical Studies*, 128(1):201–227, March 2006. ISSN 0031-8116. doi: 10.1007/s11098-005-4063-x. URL http://link.springer.com/10.1007/s11098-005-4063-x.

[174] R. Want and T. Pering. System challenges for ubiquitous & pervasive computing. In *Proceedings. 27th International Conference on Software Engineering, 2005. ICSE 2005.* Software Engineering, 2005. ICSE 2005. Proceedings. 27th International Conference on, 2005. ISBN 1-59593-963-2. doi: 10.1109/ICSE.2005.1553532.

[175] J Wayman. When bad science leads to good law: The disturbing irony of the Daubert hearing in the case of U.S V. Byron C. Mitchell. Daubert hearing on fingerprints, 2000. URL http://www.engr.sjsu.edu/biometrics/publications_daubert.html.

[176] J L Wayman. Technical testing and evaluation of biometric identification devices. In *in Biometrics: Personal Identification in Networked Society, eds. A. Jain et al. (Kluwer Academic Press*, page 345. Springer US, 1998. ISBN 978-0-387-28539-9.

[177] Zhuoshi Wei Zhuoshi Wei, Xianchao Qiu Xianchao Qiu, Zhenan Sun Zhenan Sun, and Tieniu Tan Tieniu Tan. Counterfeit iris detection based on texture analysis. In *2008 19th International Conference on Pattern Recognition*. 9th International Conference on Pattern Recognition, 2008. ISBN 978-1-4244-2174-9. doi: 10.1109/ICPR.2008.4761673.

[178] M. Weiser. The computer for the 21st Century. *IEEE Pervasive Computing*, 1(1): 94–104, 2002. ISSN 1536-1268. doi: 10.1109/MPRV.2002.993141.

[179] Richard Williams and Autonomous Systems. BAE Systems – Autonomous Capability Overview Introduction & Agenda, 2014. URL http://www.stfc.ac.uk/resources/pdf/richardwilliams.pdf.

[180] Gina Wisker. *The Postgraduate Research Handbook: Succeed with your MA, MPhil, EdD and PhD*. Palgrave Macmillan, 2nd edition, 2007. ISBN 978-0-230-52130-8. URL http://www.palgrave.com/page/detail/the-postgraduate-research-handbook-gina-wisker/.

[181] J D Woodward Jr, Christopher Horn, Julius Gatune, Aryn Thomas, and Rand Corp Santa Monica Ca. Biometrics A Look at Facial Recognition. Technical Report April, Santa Monica CA, 2005. URL http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA414520.

[182] Xiuli Xu and Ping Guo. Iris Feature Extraction Based on the Complete 2DPCA. *Advances in Neural Networks – ISNN 2009*, 5552:950–958, 2009. doi: 10.1007/978-3-642-01510-6_108. URL http://dx.doi.org/10.1007/978-3-642-01510-6_108.

[183] David Yambay, Jay Doyle, Kevin Bowyer, Adam Czajka, and Stephanie Schuckers. LivDet – Iris 2013 : Iris Liveness Detection Competition 2013. In *2013 International Conference on Biometrics (ICB)*, page 2013, 2013. ISBN 978-1-4799-3584-0. doi: 10.1109/BTAS.2014.6996283. URL http://ieeexplore.ieee.org/xpls/icp.jsp?arnumber=6613027.

[184] Roman V. Yampolskiy and Venu Govindaraju. Behavioural biometrics: a survey and classification. *International Journal of Biometrics*, 1(1):81, 2008. ISSN 1755-8301. doi: 10.1504/IJBM.2008.018665.

[185] R K Yin. *Case Study Research Design and Methods*, volume Third. 2003. ISBN 0-7619-2553-8. URL http://www.amazon.co.uk/Case-Study-Research-Methods-Applied/dp/0761925538.

[186] Y Zhang. Transparent Computing:, 2008. URL http://media.cs.tsinghua.edu.cn/~zyx/doc/transcom-euc08.pdf.

[187] Ortrun Zuber-Skerritt. Action research in higher education : examples and reflections. page 129 p., November 1996. URL http://eric.ed.gov/?id=ED351928.

Appendix A

Research Risks

Risk	Probability	Severity	Score (Prb x Sev)	Solution
Software error	1	3	3	Depending on price and time have second system set up with software and hardware as a backup.
Hardware error	1	3	3	Depending on price and time have second system set up with software and hardware as a backup.
Unobtainable Hardware/software	3	3	9	Research early the different kinds of software and hardware needed, using a product comparison and then have a three stage choice so a failure to get choice one results in getting choice two etc.
Lack of primary data	2	4	8	Begin early when gathering primary data and give people plenty of time to get into contact. Also arrange face to face meetings and supply all materials that could help the subjects.
Loss of data	2	4	8	Back up all data on multiple formats and at multiple places. Such as home/university/work/college and hard drive/external hard drive /DVD/flash drive.
Illness	2	2	4	Plan well, with plenty of time to do all tasks and time to catch-up if necessary.
Poor data	1	3	3	Collect a wide range of data and make sure that all data is relevant to the subject avoiding superfluous information.
Too much data	1	2	2	Make sure all data is catalogued in a timely fashion and concentrate on valid data.

Table A.2 Research Risks

Appendix B

Taxonomy and Algorithm Data

Liveness Measurement							
Technique	Universality	Permanence	Collectability	Performance	Acceptability	Circumvention	Total
Dermalog Algorithm	2.33	1.00	1.67	1.00	1.00	0.75	1.3
Dermalog System	2.67	0.67	5.00	3.00	1.00	1.17	2.3
Facial Modality	1.67	4.00	4.00	1.00	1.00	1.75	2.2
Feature and Texture Analysis	2.67	0.67	2.00	1.00	1.00	0.67	1.3
Static Feature Extraction	2.67	1.33	1.67	1.00	1.00	1.08	1.5
Local Binary Pattern	2.67	1.33	1.67	1.00	1.00	1.08	1.5
Texture Analysis	1.67	3.00	1.67	1.00	1.00	1.25	1.6
Purkinje Images	3.00	1.33	6.67	4.00	1.00	1.58	2.9
A V Feature Extraction	1.33	4.00	3.33	1.00	1.00	1.75	2.1
Mobile ECG	2.33	1.33	5.00	5.00	1.00	1.83	2.7

Table B.1 Overall Taxonomy Data for Liveness and Coercion Techniques

B.1 Universality Data

Liveness Measurement				
Technique	**h**	**s**	**t**	**l**
Dermalog Algorithm	1.00	3.00	3.00	1.00
Dermalog System	3.00	2.00	3.00	1.00
Facial Modality	2.00	1.00	2.00	1.00
Feature and Texture Based Analysis	3.00	2.00	3.00	1.00
Static Feature Extraction	3.00	2.00	3.00	1.00
Local Binary Pattern	3.00	2.00	3.00	1.00
Texture Analysis	1.00	1.00	3.00	1.00
Purkinje Images	4.00	2.00	3.00	1.00
Audio-Video Feature Extraction	1.00	1.00	2.00	1.00
Mobile ECG	5.00	2.00	1.00	1.00

Table B.2 Universality Liveness Calculation

Coercion Measurement				
Technique	**h**	**s**	**t**	**l**
Tangible Key Techniques	4	2	1	1
IFA	1	2	2	1
Skin Conductivity Peaks	3	1	2	1
Facial Micro-Movements	1	2	2	1

Table B.3 Universality Coercion Calculation

B.2 Permanence Data

Liveness Measurement			
Technique	**Bs**	**Ms**	**P**
Dermalog Algorithm	1.00	1.00	2.00
Dermalog System	1.00	1.00	3.00
Facial Modality	2.00	2.00	1.00
Feature and Texture Based Analysis	1.00	1.00	3.00
Static Feature Extraction	2.00	2.00	3.00
Local Binary Pattern	2.00	2.00	3.00
Texture Analysis	1.00	2.00	1.00
Purkinje Images	2.00	2.00	3.00
Audio-Video Feature Extraction	2.00	2.00	1.00
Mobile ECG	2.00	2.00	3.00

Table B.4 Permanence Liveness Calculation

Coercion Measurement			
Technique	**Bs**	**Ms**	**P**
Tangible Key Techniques	1.00	1.00	1.00
IFA	2.00	1.00	4.00
Skin Conductivity Peaks	2.00	2.00	4.00
Facial Micro-Movements	2	2.00	4.00

Table B.5 Permanence Coercion Calculation

B.3 Collectability Data

Liveness Measurement				
Technique	**Si**	**T**	**Ut**	**At**
Dermalog Algorithm	1.00	1.00	3.00	1.00
Dermalog System	1.00	1.00	3.00	3.00
Facial Modality	2.00	2.00	2.00	2.00
Feature and Texture Based Analysis	2.00	1.00	3.00	1.00
Static Feature Extraction	1.00	1.00	3.00	1.00
Local Binary Pattern	1.00	1.00	3.00	1.00
Texture Analysis	1.00	1.00	3.00	1.00
Purkinje Images	1.00	1.00	3.00	4.00
Audio-Video Feature Extraction	1.00	2.00	2.00	2.00
Mobile ECG	2.00	2.00	4.00	1.00

Table B.6 Collectability Liveness Calculation

Coercion Measurement				
Technique	**Si**	**T**	**Ut**	**At**
Tangible Key Techniques	3.00	1.00	1.00	4.00
IFA	2.00	1.00	2.00	2.00
Skin Conductivity Peaks	1.00	1.00	2.00	3.00
Facial Micro-Movements	2.00	1.00	2.00	1.00

Table B.7 Collectability Coercion Calculation

B.4 Performance Data

Liveness Measurement					
Technique	**Th**	**A**	**Ss**	**At**	**t**
Dermalog Algorithm	2.33	1.00	2.00	1.00	1.00
Dermalog System	2.67	2.00	2.00	3.00	1.00
Facial Modality	1.67	2.00	4.00	2.00	2.00
Feature and Texture Based Analysis	2.67	1.00	2.00	1.00	1.00
Static Feature Extraction	2.67	2.00	3.00	1.00	1.00
Local Binary Pattern	2.67	2.00	2.00	1.00	1.00
Texture Analysis	1.67	1.00	1.00	1.00	1.00
Purkinje Images	3.00	1.00	1.00	4.00	1.00
Audio-Video Feature Extraction	1.33	1.00	2.00	2.00	2.00
Mobile ECG	2.33	1.00	1.00	5.00	1.00

Table B.8 Performance Liveness Calculation

Coercion Measurement					
Technique	**Th**	**A**	**Ss**	**At**	**t**
Tangible Key Techniques	2.33	1.00	3.00	4.00	1.00
IFA	1.67	2.00	1.00	2.00	1.00
Skin Conductivity Peaks	2.00	2.00	1.00	3.00	1.00
Facial Micro-Movements	1.67	1.00	3.00	1.00	1.00

Table B.9 Performance Coercion Calculation

B.5 Acceptability Data

Liveness Measurement							
Technique	Vs	S	N	U	Vus	Highest Name	Metric Level
Dermalog Algorithm	5	0	0	0	0	Vs	1
Dermalog System	5	0	0	0	0	Vs	1
Facial Modality	5	0	0	0	0	Vs	1
Feature and Texture Based Analysis	5	0	0	0	0	Vs	1
Static Feature Extraction	5	0	0	0	0	Vs	1
Local Binary Pattern	5	0	0	0	0	Vs	1
Texture Analysis	5	0	0	0	0	Vs	1
Purkinje Images	5	0	0	0	0	Vs	1
Audio-Video Feature Extraction	5	0	0	0	0	Vs	1
Mobile ECG	5	0	0	0	0	Vs	1

Table B.10 Acceptability Liveness Calculation

Liveness Measurement							
Technique	Vs	S	N	U	Vus	Highest Name	Metric Level
Tangible Key Techniques	5	0	0	0	0	Vs	1
IFA	5	0	0	0	0	Vs	1
Skin Conductivity Peaks	5	0	0	0	0	Vs	1
Facial Micro-Movements	5	0	0	0	0	Vs	1
Static Feature Extraction	5	0	0	0	0	Vs	1

Table B.11 Acceptability Coercion Calculation

B.6 Circumvention Data

Liveness Measurement			
Technique	Sg	A	Ls
Dermalog Algorithm	1.00	1.00	1.00
Dermalog System	0.67	3.00	1.00
Facial Modality	4.00	1.00	2.00
Feature and Texture Based Analysis	0.67	1.00	1.00
Static Feature Extraction	1.33	1.00	2.00
Local Binary Pattern	1.33	1.00	2.00
Texture Analysis	3.00	1.00	1.00
Purkinje Images	1.33	4.00	1.00
Audio-Video Feature Extraction	4.00	1.00	2.00
Mobile ECG	1.33	5.00	1.00

Table B.12 Circumvention Liveness Calculation

Coercion Measurement			
Technique	Sg	A	Ls
Tangible Key Techniques	2.00	2.20	4.00
IFA	0.75	1.50	1.00
Skin Conductivity Peaks	1.00	1.80	1.00
Facial Micro-Movements	1.00	1.50	3.00

Table B.13 Universality Coercion Calculation

B.7 Algorithm Examples

sum t	Time Elapsed (hours)	Participants (ai/li)	Time for Response	Anomolous Userbase	ai-Ab/ai	Dr + K
1	1	5.010588235	2	1.875	0.995748299	4.162087384
3	2	6.010588235	1.64	2	0.995464853	4.162087384
5	3	7.010588235	1.65	1.625	0.996315193	4.162087384
7	4	8.010588235	1.470588235	1	0.997732426	4.162087384
9	5	9.010588235	2.25	1.25	0.997165533	4.162087384

Table B.14 Algorithm Example I

E*F	Security Level - As	Exp(ai-Ab/bi)/T	lt	ct	ai-Ab/ai
4.144391434	1.795403646	1.353374521	1.513888889	1.373722222	0.995748299
4.143211704	1.202186181	1.649988983	1.513888889	1.373722222	0.995464853
4.146750894	1.013077277	1.641384191	1.513888889	1.373722222	0.996315193
4.152649544	0.776173992	1.844244937	1.513888889	1.373722222	0.997732426
4.150290084	1.055550236	1.204705714	1.513888889	1.373722222	0.997165533

Table B.15 Algorithm Example II